

THE OPTIMAL TRILINEAR RESTRICTION ESTIMATE FOR A CLASS OF HYPERSURFACES WITH CURVATURE

IOAN BEJENARU

ABSTRACT. In [3] nearly optimal L^1 trilinear restriction estimates in \mathbb{R}^{n+1} are established under transversality assumptions only. In this paper we show that the curvature improves the range of exponents, by establishing L^p estimates, for any $p > \frac{2(n+4)}{3(n+2)}$ in the case of double-conic surfaces. The exponent $\frac{2(n+4)}{3(n+2)}$ is shown to be the universal threshold for the trilinear estimate.

1. INTRODUCTION

For $n \geq 1$, let $U \subset \mathbb{R}^n$ be an open, bounded and connected neighborhood of the origin and let $\Sigma : U \rightarrow \mathbb{R}^{n+1}$ be a smooth parametrization of an n -dimensional submanifold of \mathbb{R}^{n+1} (hypersurface), which we denote by $S = \Sigma(U)$. To this we associate the operator \mathcal{E} defined by

$$\mathcal{E}f(x) = \int_U e^{ix \cdot \Sigma(\xi)} f(\xi) d\xi.$$

Given k smooth, compact hypersurfaces $S_i \subset \mathbb{R}^{n+1}$, $i = 1, \dots, k$, where $1 \leq k \leq n+1$, the k -linear restriction estimate is the following inequality

$$(1.1) \quad \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(\mathbb{R}^{n+1})} \lesssim \Pi_{i=1}^k \|f_i\|_{L^2(U_i)}.$$

In a more compact format this estimate is abbreviated as follows:

$$\mathcal{R}^*(2 \times \dots \times 2 \rightarrow p).$$

The fundamental question regarding the above estimate is the value of the optimal p for which it holds true. Given that the estimate $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow \infty)$ is trivial, the optimality is translated into the smallest p for which the estimate holds true. In [3] Bennett, Carbery and Tao clarified the role of transversality between the surfaces involved and established that under a transversality condition between S_1, \dots, S_k the optimal exponent is $p = \frac{2}{k-1}$; the actual result in [3] is near-optimal, the optimal problem is currently open. The optimality can be easily revealed by taking S_i to be transversal hyperplanes, in which case the estimate becomes the classical Loomis-Whitney inequality.

It is also known, in some cases (precisely when $k \leq 2$), or expected, in most of the others, that curvature assumptions improve the range of exponents in (1.1), except for the case $k = n+1$. In [3], the authors state that "simple heuristics suggest that the optimal k -linear restriction theory requires at least $n+1-k$ non-vanishing principal curvatures, but that further curvature assumptions have no further effect". However, this aspect of the theory is left as an open problem in [3]. We detail below what is known for each k .

2010 *Mathematics Subject Classification.* 42B15 (Primary); 42B25 (Secondary).

Key words and phrases. Trilinear restriction estimates, Shape operator, Wave packets.

The case $k = 1$ has been understood for a very long time. Without any curvature assumptions, the optimal exponent is $p = \infty$; once the surface has some non-vanishing principal curvatures, the exponent improves to $p = \frac{2(l+2)}{l}$, where l is the number of non-vanishing principal curvatures. The case of non-zero Gaussian curvature, corresponding to $l = n$, is the classical result due to Tomas-Stein, see [19].

The case $k = 2$ without any curvature assumptions corresponds to the classical L^2 bilinear estimate, where the optimal estimate has been established. The case $k = 2$ with curvature assumptions brought a twist into the understanding of the role of curvature in this problem. The best possible exponent in $\mathcal{R}^*(2 \times 2 \rightarrow p)$ is $p = \frac{n+3}{n+1}$ and it was conjectured in [11]; it came as a surprise the fact that it was optimal not only for paraboloids (having n non-zero principal curvatures), but also for cones (having $n - 1$ non-zero principal curvatures and a vanishing principal curvature). This problem was intensely studied for most important model surfaces, see [4, 25, 22, 23, 20, 16, 17]. In [2], we addressed the problem for general surfaces and revealed a more subtle way in which the curvature comes into play by stating conditions in terms of the shape operators of the surfaces involved.

The last case for which (1.1) is fairly well-understood is the case $k = n + 1$. We note that in this case, additional curvature assumptions have no effect on the optimality of p and this is consistent with the expectation that the optimal k -linear restriction theory requires at least $n + 1 - k$ non-vanishing principal curvatures, which in this case translates into no curvature. It is conjectured that if the hypersurfaces $S_i \subset \mathbb{R}^{n+1}$ are transversal, then (1.1) holds true for $p \geq p_0 = \frac{2}{n}$. If S_i are transversal hyperplanes, (1.1) is the classical Loomis-Whitney inequality and its proof is elementary. Once the surfaces are allowed to have non-zero principal curvatures, things become far more complicated and the problem has been the subject of extensive research, see [3, 13] and references therein. In [3], Bennett, Carbery and Tao establish a near-optimal version of (1.1): this is (1.1) with an additional R^ϵ factor when the estimate is made over balls of radius R in \mathbb{R}^{n+1} . The optimal result for (1.1), that is without the ϵ -loss, is an open problem; in some cases one can use ϵ -removal techniques to derive the result without the ϵ -loss for $p > \frac{2}{n}$, see [8] for the case of surfaces with non-vanishing Gaussian curvature. The end-point for the multilinear Kakeya version of (1.1) (a slightly weaker statement than (1.1)) has been established by Guth in [13] using tools from algebraic topology.

In the remaining cases, $n \geq 3$ and $3 \leq k \leq n$, the k -linear restriction theory has been addressed in [3] where the authors established the near-optimal result for $p \geq \frac{2}{k-1}$. The exponent $\frac{2}{k-1}$ is sharp for generic surfaces, but it is not expected to be the optimal exponent once curvature assumptions are brought into the problem.

In this paper make a first step towards understanding the effect of curvature in the cases $3 \leq k \leq n$. We will argue that the correct conjectured best optimal exponent should be $p(k) = \frac{2(n+1+k)}{k(n+k-1)}$; note that this is strictly smaller than the generic optimal value $\frac{2}{k-1}$. The easy task is to show that one cannot obtain any estimates for $p < p(k)$, regardless of the amount of curvature involved; this is done by using the classical squashed-cap example. The more difficult task is to show that the conjecture is not vacuous by proving it for some class of surfaces: the multi-conical surfaces.

In this paper we look at the trilinear estimate corresponding to $k = 3$ and prove the conjecture for a particular class of surfaces: the double-conic ones. These surfaces have the nice property that they have the exact "amount" of curvature to obtain the estimate with

the optimal exponent $p(3)$, and no more, in the sense that they are "flat" in the unnecessary directions. This result is comparable to the one of Wolf [25] where he proves the bilinear restriction estimate for conical surfaces (which would translate into 1-conical surfaces in our context).

In a forthcoming paper we will provide the equivalent result for $4 \leq k \leq n$ for $k-1$ -conical surfaces. The reason to split the case $k = 3$ from $k \geq 4$ is that the argument for the latter case incurs not only additional technical difficulties, but also conceptual ones.

We now formalize the above introduction and start with the statement of the conjecture.

Conjecture. *Under appropriate transversality and curvature conditions on the surfaces S_i , $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow p)$ holds true for any $p \geq p(k) = \frac{2(n+1+k)}{k(n+k-1)}$.*

A complete resolution of this problem should have important consequences. The multilinear theory discussed above has had major impact in other problems. We mention a few such examples: In Harmonic Analysis, the bilinear and $n+1$ restriction theory was used to improve results in the context of Schrödinger maximal function, see [5, 15, 21, 10], restriction conjecture, see [23, 8, 12], the decoupling conjecture, see [7, 6]. In Partial Differential Equations, the linear theory inspired the Strichartz estimates, see [24], while the bilinear restriction theory is used in the context of more sophisticated techniques, such as the profile decomposition, see [18], and concentration compactness methods, see [14]. We expect that a positive resolution of the Conjecture will have an impact in some of the problems mentioned above.

We will make an attempt in shedding some light into the appropriate conditions required above, but we believe that only a complete resolution of the Conjecture will clarify the role of geometry into the problem.

The transversality condition is clear by now: for any choice $\zeta_i \in S_i, i = 1, \dots, k$, if $N_i(\zeta_i)$ is the unit normal to S_i , then the

$$\text{vol}(N_1(\zeta_1), \dots, N_k(\zeta_k)) \geq \nu > 0$$

where $\text{vol}(N_1(\zeta_1), \dots, N_k(\zeta_k))$ is the volume of the k -dimensional parallelepiped generated by the normals $N_i(\zeta_i)$. This is natural to impose, given that it matches the condition required in the generic case where no curvature is assumed, see [3].

The more delicate part is to list the curvature conditions on S_i . For a hypersurface $S \subset \mathbb{R}^{n+1}$ and $\zeta \in S$, we denote by $N(\zeta)$ the normal to S (assumed to be orientable) at ζ and by $S_{N(\zeta)}$ the shape operator at S at $\zeta \in S$ with respect to the normal $N(\zeta)$. For an in-depth introduction to the shape operator, we refer the reader to any classical textbook in differential geometry, see for instance [9]. For a quick introduction tailored towards the use in the context of bilinear estimates, see [2]. Note that we use the same letter S both for the hypersurface and the shape operator: the subscript N will indicate that we are referring to the shape operator.

As for the role of curvature in the above Conjecture, we anticipate a condition of type

$$\text{vol}(N_1(\zeta_1), \dots, N_k(\zeta_k), S_{N_j(\zeta_j)}v_{k+1}, \dots, S_{N_j(\zeta_j)}v_{n+1}) \geq \nu > 0$$

for any choices $\zeta_i \in S_i, i = 1, \dots, k$, for any $j \in \{1, \dots, k\}$, for any choice $\zeta_j \in S_j$ and for any choice v_{k+1}, \dots, v_{n+1} is a orthonormal vector basis to specific $n-k+1$ -dimensional submanifolds $S \subset S_j$ (it suffices to check this for one basis). Here $S_{N_i(\zeta_i)}$ is the shape operator at S_i at

$\zeta_i \in S_i$ with respect to the normal $N_i(\zeta_i)$. We expect these submanifolds to be exhausted by the family

$$S(\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_{n+1}) = S_j \cap_{l \neq j} (\beta_l + S_l).$$

for any choices $\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_{n+1} \in \mathbb{R}^{n+1}$. But we do not go as far as claiming this to be the ultimate curvature condition.

In Section 7 we show, by a counterexample, that the range stated in the conjecture is sharp, that is (1.1) fails for $p < \frac{2(n+1+k)}{k(n+k-1)}$ for conic surfaces, but also for quadratic surfaces. This highlights that regardless of how much curvature is brought into the problem, the threshold $\frac{2(n+1+k)}{k(n+k-1)}$ is the best exponent that can occur in $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow p)$.

As for positive results, in this paper we look at the case $k = 3$ and establish the almost optimal conjectured result for a particular class of hypersurfaces which we introduce below.

We recall the definition of a foliation. A 2-dimensional foliation of the (n -dimensional) hypersurface S is a decomposition of S into a union of connected disjoint sets $\{S_\alpha\}_{\alpha \in A}$, called the leaves of the foliation, with the following property: every point in S has a neighborhood V and local system of coordinates $x : V \subset S \rightarrow \mathbb{R}^n$ such that for each leaf S_α , the coordinates of $V \cap S_\alpha$ are $\xi_3 = \text{constant}, \dots, \xi_n = \text{constant}$.

We are now ready to formalize the conditions we impose on our surfaces. As before, $S_i, i \in \{1, 2, 3\}$ are hypersurfaces with smooth parameterizations $\Sigma_i : U_i \subset \mathbb{R}^n \rightarrow \mathbb{R}^{n+1}$, where each U_i are open, bounded and connected neighborhood of the origin (note that different U_i may belong to different hyperplanes identified with the same \mathbb{R}^n). In addition, we assume the following three hypothesis:

- i) (foliation) for each $i \in \{1, 2, 3\}$, the hypersurface S_i admits the foliation

$$S_i = \bigcup_{\alpha} S_{i,\alpha}$$

where, for each α , the leaf $S_{i,\alpha}$ is a flat submanifold of dimension 2.

- ii) (the leaves are completely flat) If $S_{N_i(\zeta_i)}$ is the shape operator of S_i at $\zeta_i \in S_i$ with choice of normal $N_i(\zeta_i)$ we assume that for every $v \in T_{\zeta_i} S_{i,\alpha}$ (the tangent plane at $S_{i,\alpha}$ at the point $\zeta_i \in S_{i,\alpha}$) the following holds

$$S_{N_i(\zeta_i)} v = 0.$$

- iii) (transversality and curvature) For any $\zeta_i \in S_i, i \in \{1, 2, 3\}$, for any $l \in \{1, 2, 3\}$ and for any orthonormal basis $v_4, \dots, v_{n+1} \in (T_{\zeta_i} S_{l,\alpha})^\perp \subset T_{\zeta_i} S_{l,\alpha}$ the following holds true

$$(1.2) \quad \text{vol}(N_1(\zeta_1), N_2(\zeta_2), N_3(\zeta_3), S_{N_l(\zeta_l)} v_4, \dots, S_{N_l(\zeta_l)} v_{n+1}) \geq \nu.$$

In (1.2) the vol is the standard volume form of $n+1$ vectors in \mathbb{R}^{n+1} , thus the condition quantifies the linear independence of the vectors $N_1(\zeta_1), N_2(\zeta_2), N_3(\zeta_3), S_{N_l(\zeta_l)} v_4, \dots, S_{N_l(\zeta_l)} v_{n+1}$.

The condition ii) says that $S_{i,\alpha}$ are, in some sense, completely flat components of S_i since, besides being subsets of affine planes of dimension 2, the normal $N(\zeta)$ to S_i is constant as we vary ζ along $S_{i,\alpha}$ for fixed α .

The first things to read in condition iii) is the transversality condition between S_1, S_2 and S_3 due to the transversality between any choice on normals. The condition iii) also says that the submanifolds transversal to the leafs carry the curvature assumptions, in the sense that their tangent space does not contain any eigenvectors of the shape operator. In addition, for each $i \in \{1, 2, 3\}$, we are guaranteed to have transversality between $N_1(\zeta_1), N_2(\zeta_2), N_3(\zeta_3)$ and $S_{N_i} T_{\zeta_i}(S_{l,\alpha})^\perp$.

At this point we can state the main result of this paper.

Theorem 1.1. *Assume that S_1, S_2, S_3 satisfy the conditions i)-iii) above. Given any p with $p(3) = \frac{2(n+4)}{3(n+2)} < p \leq \infty$, the following holds true*

$$(1.3) \quad \|\Pi_{i=1}^3 \mathcal{E}_i f_i\|_{L^p(\mathbb{R}^{n+1})} \leq C(p) \Pi_{i=1}^3 \|f_i\|_{L^2(U_i)}, \quad \forall f_i \in L^2(U_i).$$

To our best knowledge this result is the first instance when the trilinear restriction estimate is shown to improve its range of exponents when curvature assumptions are made; this statement has to be carefully explained, since given large enough n one can obtain trilinear restriction estimates in L^p with $p < 1$ based on linear and bilinear restriction estimates, for the latter one see [17]. We are not only going beyond the known L^1 estimate, but also provide the sharp result, except for the endpoint $p = p(3)$, and this is truly a feature of the trilinear estimate. In some instances, our result also improves the known results in the range $p \geq 1$, where the standard result is near-optimal, while ours is optimal. Given that the estimate is trivial for $p = \infty$, it suffices to focus on the result above in the cases $p(3) < p \leq 1$ and this is what we will do.

One aspect that is revealed by the above result is that the optimal trilinear restriction estimate discards the effect of 2 curvatures; indeed, each S_i has precisely 2 vanishing principal curvatures. In light of the similar results in the bilinear case, this result indicates that the optimal k -linear restriction estimate discards the effect of $k - 1$ curvatures and relies only on $n + 1 - k$ curvatures, although the actual statements have to be more rigorous.

Now we list some hypersurfaces that would fit in our class and for each the above result is applicable. One such model is given by the equation:

$$(1.4) \quad \zeta_{n+1} + \zeta_n = \sqrt{\zeta_1^2 + \dots + \zeta_{n-1}^2}.$$

Away from the origin, this surface has 2 vanishing principal curvatures and $n - 2$ non-vanishing principal curvatures. If one writes $(\zeta_1, \dots, \zeta_{n-1}) = r \cdot \omega$ with $r = \sqrt{\zeta_1^2 + \dots + \zeta_{n-1}^2}$ and $\omega \in \mathbb{S}^{n-2}$ the above equation becomes $\zeta_{n+1} + \zeta_n = r$, thus revealing the source of the vanishing curvatures as coming from the embedded planes of dimension 2; these are the leaves of the foliation.

Another model is given by the cylinders:

$$\zeta_1^2 + \dots + \zeta_{n-1}^2 = 1$$

where the freedom in choosing ζ_n and ζ_{n+1} gives rise to the desired structure.

At this time we are not able to establish a similar result for more general surfaces, in particular for hypersurfaces with non-zero Gaussian curvature, such as the paraboloid $\zeta_{n+1} = \zeta_1^2 + \dots + \zeta_n^2$. This may come as a surprise as one expects more curvature to be helpful. By now there is enough body of evidence that additional curvature complicates these problems. Probably the most clear piece of evidence lies in the $n + 1$ -linear estimate in R^{n+1} (where no curvature is needed for optimal results): if the surfaces are transversal hyperplanes, the $L^{\frac{1}{n}}$ estimate follows easily from the classical Loomis-Whitney inequality, while if the surfaces are assumed to have curvature, the estimate is considerably harder and open.

We now explain some of the key ideas in the proof of Theorem 1.1. Some of the starting ideas originate in the prior work of the author on the bilinear restriction estimate in [2] and the multilinear restriction estimate in [1].

The argument for the bilinear restriction estimate that we provided in [2] was inspired by many previous works in the subject, and we chose to follow the less common approach of Tao in [22]. One of the novelties of our argument is that it reveals in a clear way a key feature of the argument: at some stage of it, the proof relies on replicating the argument of the classical bilinear L^2 estimate (on the Fourier side, thanks to the Plancherel formula). To be more specific, inside the body of the argument, we reproduce a proof of the bilinear estimate

$$(1.5) \quad \|\mathcal{E}_1 f_1 \cdot \mathcal{E}_2 f_2\|_{L^2} \lesssim \|f_1\|_{L^2(U_1)} \|f_2\|_{L^2(U_2)},$$

where transversality assumptions are made on $S_1 = \Sigma_1(U_1)$ and $S_2 = \Sigma_2(U_2)$. This bilinear estimate is used in a localized setup, meaning that one quantifies the effect of U_1 or U_2 having small support. We note that in the combinatorial strategy of proving the bilinear restriction estimate, used for instance in [23], one replaces the bilinear L^2 estimate with its quadrilinear version and the relevance of the bilinear L^2 estimate (1.5) is less apparent.

Thus the natural idea is to follow the same route in the trilinear setup, that is to rely, at some stage of the argument, on the trilinear estimate

$$(1.6) \quad \|\mathcal{E}_1 f_1 \cdot \mathcal{E}_2 f_2 \cdot \mathcal{E}_3 f_3\|_{L^1} \lesssim \|f_1\|_{L^2(U_1)} \|f_2\|_{L^2(U_2)} \|f_3\|_{L^2(U_3)}.$$

This is where one hits an immediate road-block. While the bilinear L^2 estimate is well understood and very malleable, with arguments that easily blend into those used in [2], we found the trilinear L^1 estimate to be very rigid. In particular, one needs to use the estimate in a black-box fashion. Moreover, one needs to use a localized version of the trilinear estimate, in the sense that one of the factors has small frequency support, or, equivalently, the surface has small support in some directions. We have not found such a result in literature and probably for a good reason: all the previous proofs for the multilinear restriction estimates seem to go through the multilinear Keakeya and it is not clear how to translate the localized multilinear estimate into a multilinear Keakeya estimate. This was part of the motivation for our work in [1] where we provided a direct proof of the multilinear estimate, followed by a refinement in the context of small support. In Section 5 we adapt the result from [1] to our current needs.

The next tool is to construct tables adapted to waves which are able to highlight the dispersive effects in outer regions. This is done by using a wave packet decomposition and as in [2], we use of the refined wave packet decomposition introduced by Tao in [22]. This requires a bit more work than the classical wave packet construction, but it has the advantage of providing a direct venue for closing the argument for $p > p(3)$ without additional ϵ -removal techniques, see [20] for example. This is important because the ϵ -removal results, see [20, Lemma 2.4], work for L^p spaces with $p > 1$, hence they do not cover the ranges of L^p spaces we are interested in.

We expect to be able to obtain a similar result to the one in Theorem 1.1 for the k -linear restriction estimate with $k \geq 4$. The argument provided here fails to work for $k \geq 4$ and we can easily point to such an instance: the key estimate (6.4) is an L^1 estimate and the triangle inequality holds in this space, but it fails in $L^{\frac{2}{k-1}}$ (and this is the space to be used in the corresponding step for the k -linear restriction estimate). But, in fact, there are more subtle points where the argument here simply fails to carry over to the case $k \geq 4$.

1.1. Notation. We start by clarifying the role of various constants that appear in the argument. N is a large integer that depends only on the dimension. C is a large constant that may change from line to line, may depend on N , but not on c and C_0 introduced below. C is used in the definition of: $A \lesssim B$, meaning $A \leq CB$, $A \ll B$, meaning $A \leq C^{-1}B$, and $A \approx B$, meaning $A \lesssim B \wedge B \lesssim A$. For a given number $r \geq 0$, by $A = O(r)$ we mean that $A \approx r$. C_0 is a constant that is independent of any other constant and its role is to reduce the size of cubes in the inductive argument. It can be set $C_0 = 4$ throughout the argument, but we keep it this way so that its role in the argument is not lost. $c \ll 1$ is a very small variable meant to make expressions $\ll 1$ and most estimates will be stated to hold in a range of c .

We use the standard notation $(\xi_1, \dots, \bar{\xi}_i, \dots, \xi_l) := (\xi_1, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_l)$ to mean that one component is missing.

By powers of type $R^{\alpha+}$ we mean $R^{\alpha+\epsilon}$ for arbitrary $\epsilon > 0$. Practically they should be seen as $R^{\alpha+\epsilon}$ for arbitrary $0 < \epsilon \lesssim 1$. The estimates where such powers occur will obviously depend on ϵ .

By $B(x, R)$ we denote the ball centered at x with radius R in the underlying space (most of the time it will be \mathbb{R}^n or \mathbb{R}^{n+1}).

Let $\eta_0 : \mathbb{R}^n \rightarrow [0, +\infty)$ be a Schwartz function, normalized in L^1 , that is $\|\eta_0\|_{L^1} = 1$, and with Fourier transform supported on the unit ball. Given some $r > 0$ we denote by $\eta_r(x) = r^{-n}\eta_0(r^{-1}x)$ and note that $\hat{\eta}_r$ is supported in $B(0, r)$.

A disk $D \subset \mathbb{R}^{n+1}$ has the form

$$D = D(x_D, t_D; r_D) = \{(x, t_D) \in \mathbb{R}^{n+1} : |x - x_D| \leq r_D\},$$

for some $(x_D, t_D) \in \mathbb{R}^{n+1}$ and $r_D > 0$. We define the associated smooth cut-off

$$\tilde{\chi}_D(x, t) = (1 + \frac{|x - x_D|}{r_D})^{-N}.$$

A cube $Q \subset \mathbb{R}^{n+1}$ of size R has the standard definition $Q = \{(x, t) \in \mathbb{R}^{n+1} : \|(x - x_Q, t - t_Q)\|_{l^\infty} \leq \frac{R}{2}\}$, where $c_Q = (x_Q, t_Q)$ is the center of the cube. Given a constant $\alpha > 0$ we define αQ to be the dilated by α of Q from its center, that is $\alpha Q = \{(x, t) \in \mathbb{R}^{n+1} : \|(x - x_Q, t - t_Q)\|_{l^\infty} \leq \alpha \cdot \frac{R}{2}\}$.

Given a cube $q \subset \mathbb{R}^{n+1}$ of size r we will use three functions that are highly concentrated in q . The classic one is the characteristic function χ_q^c of q . We reserve the more standard notation χ_q for the next object which is used more frequently. This is build with the help of η_0 (we abuse notation here as we should be using the corresponding $\eta_0 : \mathbb{R}^{n+1} \rightarrow [0, +\infty)$ with similar properties):

$$\chi_q(x) = \eta_0\left(\frac{x - c(q)}{r}\right).$$

This localization function has nice properties on the Fourier side. The other localization function is

$$\tilde{\chi}_q(x) = (1 + |\frac{x - c(q)}{r}|)^{-N},$$

for some large N . This localization has better properties on the physical side.

Acknowledgement. Part of this work was supported by a grant from the Simons Foundation (#359929, Ioan Bejenaru). Part of this work was supported by the National Science Foundation under grant No. 0932078000 while the author was in residence at the Mathematical Research Sciences Institute in Berkeley, California, during the Fall 2015 semester.

2. GEOMETRY OF THE SURFACES AND CONSEQUENCES

We start by simplifying the setup. The surfaces are bounded, therefore we can always break them into smaller (and similar) pieces such as to accommodate the additional hypothesis described below.

First we note that we can assume each S_i to be of graph type: there is a smooth map $\varphi_i : U_i \subset \mathbb{R}^n \rightarrow \mathbb{R}$ such that $S = \{\Sigma_i(\xi) = (\xi, \varphi_i(\xi)) : \xi \in U_i\}$. Here U_i are open, connected with compact closure. It is less important that the graphs are of type $\zeta_{n+1} = \varphi_i(\zeta_1, \dots, \zeta_n)$ (we can have as well $\zeta_k = \varphi_i(\zeta_1, \dots, \bar{\zeta}_k, \dots, \zeta_{n+1})$), although we can accommodate this by a rotation of coordinates. Then each flat leaf $S_{i,\alpha}$ corresponds to a flat leaf $U_{i,\alpha}$, in the sense that $\Sigma_i(U_{i,\alpha}) = S_{i,\alpha}$; this is indeed the case since projections onto hyperplanes along a vector transversal to S_i takes 2-dimensional affine planes to 2-dimensional affine planes.

We can find a system of coordinates $\mathbf{x}_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ that parametrizes each leaf $U_{i,\alpha}$ into a new flat leaf $\tilde{U}_{i,\alpha}$ characterized by $\xi_3 = \text{constant}, \dots, \xi_n = \text{constant}$. Finally, we assume that each U_i has small enough diameter.

Next, we derive a key geometric consequence of our setup. Given a surface S_i we define $\mathcal{N}_i = \{N_i(\zeta_i) : \zeta_i \in S_i\}$ be the set of normals at S_i . By $dspan\mathcal{N}_i$ we denote the following subset of the classical span of \mathcal{N}_i :

$$dspan\mathcal{N}_i = \{\alpha N_\alpha + \beta N_\beta : N_\alpha, N_\beta \in \mathcal{N}_i, \alpha, \beta \in \mathbb{R}\}.$$

$dspan\mathcal{N}_i$ is the set of linear combinations of two vectors in \mathcal{N}_i ; it is not a linear subspace. With these notation in place, we claim the following result.

Lemma 2.1. *If S_1, S_2, S_3 satisfy the conditions i)-iii), then for any $N \in dspan\mathcal{N}_1$ and any $N_2 \in \mathcal{N}_2, N_3 \in \mathcal{N}_3$ the following holds true for all real numbers a, b, c :*

$$(2.1) \quad |aN + bN_2 + cN_3| \gtrsim \max(|a||N|, |b|, |c|).$$

The statement is symmetric with respect to S_1, S_2, S_3 .

We note that the occurrence of $|N|$ in $\max(|a||N|, |b|, |c|)$ is motivated by the fact that vectors in $dspan\mathcal{N}_1$ are not normalized, but vectors in \mathcal{N}_2 and \mathcal{N}_3 are.

Proof. (2.1) is equivalent to a transversality condition between N, N_2, N_3 :

$$vol(N, N_2, N_3) \gtrsim |N|$$

We write $N = \alpha N_\alpha + \beta N_\beta$ for some $N_\alpha \neq N_\beta$ and consider $\gamma : [0, t_0] \rightarrow S_1$, a smooth curve with the property that $N_1(\gamma(0)) = N_\alpha$ and $N_1(\gamma(t_0)) = N_\beta$. We also assume that $|\gamma'(t)| = 1$ on $[0, t_0]$ and that $0 \leq t_0 \ll 1$; this is possible because we assumed U_1 of small diameter. In addition, if α_0 is such that $\gamma(0) \in S_{i,\alpha_0}$, we can assume that $\gamma'(0) \in (T_{\gamma(0)}S_{1,\alpha_0})^\perp$. Then we have

$$\begin{aligned} N_1(\gamma(t_0)) &= N_1(\gamma(0)) + \int_0^{t_0} S_{N_1(\gamma(s))} \gamma'(s) ds \\ &= N_1(\gamma(0)) + t_0 S_{N_1(\gamma(0))} \gamma'(0) + O(t_0^2). \end{aligned}$$

We then continue with

$$\begin{aligned} N &= \alpha N_1(\gamma(0)) + \beta(N_1 + t_0 S_{N_1(\gamma(0))} \gamma'(0) + O(t_0^2)) \\ &= (\alpha + \beta) N_1(\gamma(0)) + \beta t_0 S_{N_1(\gamma(0))} \gamma'(0) + \beta O(t_0^2) \end{aligned}$$

The two vectors $N_1(\gamma(0))$ and $S_{N_1(\gamma(0))} \gamma'(0)$ are transversal, thus $|N| \approx |\alpha + \beta| + t_0 |\beta| |S_{N_1(\gamma(0))} \gamma'(0)|$ (here we use that $t_0 \ll 1$), and also

$$\begin{aligned} \text{vol}(N, N_2, N_3) &\approx \text{vol}((\alpha + \beta) N_1(\gamma(0)) + \beta t_0 S_{N_1(\gamma(0))} \gamma'(0), N_2, N_3) \\ &\gtrsim |(\alpha + \beta) N_1(\gamma(0)) + \beta t_0 S_{N_1(\gamma(0))} \gamma'(0)| \approx |N| \end{aligned}$$

where we have used the following consequence of (1.2):

$$\text{vol}(N_1(\gamma(0)), S_{N_1(\gamma(0))} v, N_2, N_3) \gtrsim 1,$$

which holds true for any unit vector $v \in (T_{\gamma(0)} S_{1, \alpha_0})^\perp \subset T_{\gamma(0)} S_1$. \square

Using a similar argument as above, one can easily establish the following dispersive estimate

$$(2.2) \quad |N_i(\zeta_1) - N_i(\zeta_2)| \approx d(S_{i, \alpha_1}, S_{i, \alpha_2})$$

where $S_{i, \alpha_1}, S_{i, \alpha_2}$ are the leafs to which ζ_1, ζ_2 belong to, respectively. Here the distance between S_{i, α_1} and S_{i, α_2} can be defined either by using geodesics inside the hypersurface S_i (using the induced metric from the ambient space R^{n+1}) or, equivalently, by using the classical distance between sets in R^{n+1} .

3. RESTATING THE PROBLEM

3.1. Rephrasing the problem in terms of free waves. We reformulate our problem in terms of free waves, this being motivated by the use of wave packets in order to prove Theorem 1.1. Once the wave packet decomposition is made and its properties of the packets are clear, the formalization of the problem as an evolution equation can be forgotten.

Assume we are given a surface S with a graph type parametrization $\zeta_{n+1} = \varphi(\xi)$ where $\xi = (\zeta_1, \dots, \zeta_n)$. We rename the variable ζ_{n+1} by τ , thus the equation of S becomes $\tau = \varphi(\xi)$. We parametrize the physical space by $(x, t) \in \mathbb{R}^n \times \mathbb{R}$. We make the choice that τ is the Fourier variable corresponding to t , while ξ is the Fourier variable corresponding to x . In what follows we use the convention that \hat{f} denotes the Fourier transform of f with respect to the x variable.

We define the free wave $\phi = \mathcal{E}f$ as follows

$$\phi(x, t) = \mathcal{E}f(x, t) = \int_{\mathbb{R}^n} e^{i(x \cdot \xi + t \varphi(\xi))} f(\xi) d\xi.$$

Note that $\phi(0) = \check{f}$ and $\hat{\phi}(\xi, t) = e^{it\varphi(\xi)} \hat{\phi}(\xi, 0)$. We define the mass of a free wave by $M(\phi(t)) := \|\phi(t)\|_{L^2}^2$ and note that it is time independent:

$$M(\phi(t)) := \|\phi(t)\|_{L^2}^2 = \|\hat{\phi}(t)\|_{L^2}^2 = \|\hat{\phi}(0)\|_{L^2}^2 = \|\phi(0)\|_{L^2}^2 = M(\phi(0)).$$

The proof of (1.3) relies on estimating $\Pi_{i=1}^3 \mathcal{E}_i f_i$ on cubes on the physical side and see how this behaves as the size of the cube goes to infinity by using an inductive type argument with respect to the size of the cube. Before we formalize this strategy, we note that at every stage of the inductive argument we re-localize functions both on the physical and frequency

space, and, as a consequence, we need to quantify the new support on the frequency side. This will be done by using the margin of a function.

We assume we are given a reference set V inside which we want to keep all functions supported. If f is supported in $U \subset V$ we define the margin of f relative to V by

$$\text{margin}(f) := \text{dist}(\text{supp}(f), V^c).$$

In terms of free waves $\phi = \mathcal{E}f$, the margin is defined by

$$\text{margin}(\phi(t)) := \text{dist}(\text{supp}_\xi(\hat{\phi}(t)), V^c) = \text{dist}(\text{supp}(f), V^c),$$

where we have used that the Fourier support of $\hat{\phi}(t)$ is time independent and that $\hat{\phi}(0) = f$. In other words, the margin of a free wave is time independent.

In practice, we work with three different types of free waves, $\phi_i = \mathcal{E}_i f_i$. They are assumed to be graphs with different phase functions φ_i and with potentially different ambient domain, that is U_i are subsets of different subspaces isomorphic to \mathbb{R}^n (for instance U_i can be subsets of the hyperplanes $\xi_i = 0$). The above construction changes only in making the choice of τ being the coordinate in the direction normal to the ambient hyperplane to which U_i belongs to, while ξ are the coordinates in the ambient hyperplane. Obviously, the margin of each ϕ_i is then defined with respect to some V_i into the same ambient hyperplane. When choosing the reference sets V_i we need to impose that the conditions i)-iii) hold true on $\Sigma_i(V_i)$ as well.

Next, we prepare the elements that are needed for the induction on scale argument.

Definition 3.1. *Let $p(3) \leq p \leq 1$. Given $R \geq C_0$ we define $A_p(R)$ to be the best constant for which the estimate*

$$(3.1) \quad \|\Pi_{i=1}^k \phi_i\|_{L^p(Q_R)} \leq A_p(R) \Pi_{i=1}^k M(\phi_i)^{\frac{1}{2}}.$$

holds true for all cubes Q_R of size-length R , $\phi_i = \mathcal{E}_i f_i$ and obeying the margin requirement

$$(3.2) \quad \text{margin}^i(\phi_i) \geq M - R^{-\frac{1}{4}}, i = 1, \dots, k.$$

The goal is to obtain an uniform estimate on $A_p(R)$ with respect to R . In the absence of the margin requirement above, $A_p(R)$ would be an increasing function. However, since the argument needs to tolerate the margin relaxation, we also define

$$\bar{A}_p(R) := \sup_{1 \leq r \leq R} A_p(r)$$

and the new $\bar{A}_p(R)$ is obviously increasing with respect to R .

Then (1.3), and as a consequence the main result of this paper, Theorem 1.1, follow from the next result.

Proposition 3.2. *Assume $0 < \epsilon < 1$. If $R \gg 2^{2C_0}$ and $R^{-\frac{1}{4}+} \ll c \ll 1$, there exists $C(\epsilon)$ such that the following holds true:*

$$(3.3) \quad A_p(R) \leq (1 + cC) \left((1 + cC)^p \left(\bar{A}_p\left(\frac{R}{2}\right) \right)^p + \left(C(\epsilon) c^{-C} R^{\frac{n+4}{2}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4}) + \epsilon} \right)^p \right)^{\frac{1}{p}}.$$

We show how (1.3) follows from (3.3). From the beginning we can assume the $\bar{A}_p(R) \geq 1$; from this we obtain the simpler inequality

$$A_p(R) \leq (1 + cC)^2 \bar{A}_p\left(\frac{R}{2}\right) \left(1 + \left(C(\epsilon) c^{-C} R^{\frac{n+4}{2}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4}) + \epsilon} \right)^p \right)^{\frac{1}{p}}.$$

Since $p > \frac{2}{3} \cdot \frac{n+4}{n+2}$, we set $c^{-C} = R^{-\frac{n+4}{4}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4}) - \epsilon}$, that is $c = R^{\frac{n+4}{4C}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4}) + \frac{\epsilon}{C}}$, and note that c satisfies $R^{-\frac{1}{4}+} \ll c \ll 1$, provided $C(n, p)$ is large enough and ϵ is chosen small enough such that $\frac{n+4}{4}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4}) + \epsilon < 0$. Then we apply the previous inequality to obtain

$$A_p(R) \leq (1 + cC)^2 \bar{A}_p\left(\frac{R}{2}\right) \left(1 + \left(C(\epsilon) R^{\frac{n+4}{4}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4})}\right)^p\right)^{\frac{1}{p}}.$$

For R large enough, $C(\epsilon) R^{\frac{n+4}{4}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4})} < 1$, and by applying Bernoulli's inequality, $(1+x)^r \leq 1 + (2^r - 1)x$ for $|x| \leq 1$ and $0 < r < 1$, we continue with

$$A_p(R) \leq (1 + cC)^2 \bar{A}_p\left(\frac{R}{2}\right) \left(1 + (2^{\frac{1}{p}} - 1)C(\epsilon) R^{\frac{n+4}{4}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4})}\right).$$

Taking the maximum with respect to $r \in [\frac{R}{2}, R]$ gives

$$\bar{A}_p(R) \leq (1 + CR^{\frac{n+4}{4C}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4})})^2 \bar{A}_p\left(\frac{R}{2}\right) \left(1 + (2^{\frac{1}{p}} - 1)C(\epsilon) R^{\frac{n+4}{4}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4})}\right).$$

Since both powers of R are negative, $\frac{n+3}{4C}(\frac{1}{p} - \frac{n+1}{n+3})$, $\frac{n+3}{4}(\frac{1}{p} - \frac{n+1}{n+3}) < 0$, this estimate can be iterated to show that $\bar{A}_p(R)$ is uniformly bounded in terms of $\bar{A}_p(C^\alpha 2^{2C_0})$ for all $R \geq C^\alpha 2^{2C_0}$; here α is chose large enough so that if $R \geq C^\alpha 2^{2C_0}$, then $C(\epsilon) R^{\frac{n+4}{4}(\frac{1}{p} - \frac{3}{2} \cdot \frac{n+2}{n+4})} < 1$. Since $\bar{A}_p(C^\alpha 2^{2C_0})$ is bounded by a constant depending on C_0 , C and p (and p influences the choice of ϵ), (1.3) follows and we conclude the proof of Theorem 1.1.

3.2. Tables on cubes. Let $Q \subset \mathbb{R}^{n+1}$ be a cube of radius R . Given $j \in \mathbb{N}$ we split Q into $2^{(n+1)j}$ cubes of size $2^{-j}R$ and denote this family by $\mathcal{Q}_j(Q)$; thus we have $Q = \cup_{q \in \mathcal{Q}_j(Q)} q$. If $j \in \mathbb{N}$ and $0 \leq c \ll 1$ we define the (c, j) interior $I^{c,j}(Q)$ of Q by

$$(3.4) \quad I^{c,j}(Q) := \bigcup_{q \in \mathcal{Q}_j(Q)} (1 - c)q.$$

Given $j \in \mathbb{N}$ we define a table Φ on Q to be a vector $\Phi = (\Phi^{(q)})_{q \in \mathcal{Q}_j(Q)}$ and define its mass by

$$M(\Phi) = \sum_{q \in \mathcal{Q}_j(Q)} M(\Phi^{(q)}).$$

We define the margin of a table as the minimum margin of its components:

$$\text{margin}(\Phi) = \min_{q \in \mathcal{Q}_j(Q)} \text{margin}(\Phi^{(q)}).$$

Inspired by the Lemma 6.1 in [22], we will make use of the following result.

Lemma 3.3. *Assume $0 < p < \infty$, $R \gg 1$, $0 < c \ll 1$ and f smooth. Given a cube $Q_R \subset \mathbb{R}^{n+1}$ of size R , there exists a cube Q of size $2R$ contained in $4Q_R$ such that*

$$(3.5) \quad \|f\|_{L^p(Q_R)} \leq (1 + cC) \|f\|_{L^p(I^{c,j}(Q))}$$

Proof. Using Fubini's theorem, we have the following identity

$$\int_{Q_R} \|f\|_{L^p((Q_R \cap I^{c,j}(Q(x,t;2R)))}^p dx dt = \int_{Q_R} |f(x,t)|^p |Q_R \cap I^{c,j}(Q(x,t;2R))| dx dt.$$

From the definition of $I^{c,j}(Q(x,t;2R))$ it follows that

$$|Q(x,t;2R) \setminus I^{c,j}(Q(x,t;2R))| \leq (n+1)c|Q(x,t;2R)| = (n+1)2^{n+1}c|Q_R|$$

and, as a consequence,

$$|Q_R| \leq (1 + (n+1)2^{n+1}c)|Q_R \cap I^{c,j}(Q(x,t;2R))|, \quad \forall (x,t) \in Q_R.$$

In the above we have used that if $(x,t) \in Q_R$ then $Q_R \subset Q(x,t;2R)$.

Combining this estimates with the above identity, leads to

$$\|f\|_{L^p}^p \leq \frac{1}{|Q_R|} \int_{Q_R} (1 + (n+1)2^{n+1}c) \|f\|_{L^p((Q_R \cap I^{c,j}(Q(x,t;2R)))}^p dx dt$$

By the pigeonholing principle, it follows that there is $(x,t) \in Q_R$ such that

$$\|f\|_{L^p}^p \leq (1 + (n+1)2^{n+1}c) \|f\|_{L^p((Q_R \cap I^{c,j}(Q(x,t;2R)))}^p$$

and since $(1 + (n+1)2^{n+1}c)^{\frac{1}{p}} \leq 1 + cC$, the conclusion follows. Note that here the value of p affects the choice of C . \square

4. WAVE PACKETS

The wave packet construction is standard by now. The key elements are: an interpretation of the $\phi = \mathcal{E}f$ as an evolution equation with initial data $\phi(0) = \check{f}$, a phase-spaces decomposition of the initial data in a linear space and an analysis of the evolution of each such component. The need of all these elements is the reason to require the more restrictive hypothesis of graph-like foliation.

We now continue with the formalization of the wave packet construction for double-conical surfaces. We assume that S is of double-conic type and it has the graph-type parametrization $\Sigma : U \rightarrow S$, where $\Sigma(\xi) = (\xi, \varphi(\xi))$ and with foliations $U = \cup_{\alpha} U_{\alpha}$, $S = \cup_{\alpha} S_{\alpha}$, $\Sigma(U_{\alpha}) = S_{\alpha}$.

For the foliation $U = \cup_{\alpha} U_{\alpha}$, we choose a system of coordinates $\mathbf{x} : U \rightarrow \tilde{U}$ such that for each leaf U_{α} , the coordinates of U_{α} are $\xi_3 = \text{constant}, \dots, \xi_n = \text{constant}$. Let $\tilde{U}' = \pi(\tilde{U})$, where $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^{n-2}$ is the projection $\pi(\xi_1, \dots, \xi_n) = (\xi_3, \dots, \xi_n)$. Let $\tilde{\mathcal{L}}$ be a maximal r^{-1} -separated subset of $\tilde{U}' \subset \mathbb{R}^{n-2}$. For each $\tilde{\xi} \in \tilde{\mathcal{L}}$, $\mathbf{x}^{-1}(\cdot, \cdot, \tilde{\xi})$ is a leaf, that is $\mathbf{x}^{-1}(\cdot, \cdot, \tilde{\xi}) = U_{\alpha}$ for some α . In each such leaf we pick ξ_T and call \mathcal{L} to be the set obtained this way. It is not important which $\xi_T \in \mathbf{x}^{-1}(\cdot, \cdot, \tilde{\xi})$ is chosen, since from condition ii) it follows that, for $\xi \in U_{\alpha}$, the normal $N(\Sigma(\xi))$ to S is constant as ξ varies inside the leaf U_{α} . We denote by $U(\xi_T)$ the leaf U_{α} to which ξ_T belongs and by $S(\xi_T) = \Sigma(U(\xi_T))$, the corresponding leaf on S . We note that $d(U(\xi_{T_1}), U(\xi_{T_2})) \approx d(\tilde{\xi}_1, \tilde{\xi}_2)$ which combined with (2.2) gives

$$(4.1) \quad |N(\Sigma(\xi_{T_1})) - N(\Sigma(\xi_{T_2}))| \approx d(U(\xi_{T_1}), U(\xi_{T_2})) \approx d(\tilde{\xi}_1, \tilde{\xi}_2).$$

Let L be the lattice $L = c^{-2}r\mathbb{Z}^n$. With $x_T \in L$, $\xi_T \in \mathcal{L}$ we define the tube $T = T(x_T, \xi_T) := \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : |x - x_T + t\nabla\varphi(\xi_T)| \leq c^{-2}r\}$ and denote by \mathcal{T} the set of such tubes. One notices that T is the $c^{-2}r$ neighborhood of the line passing through $(x_T, 0)$ and direction $N(\Sigma(\xi_T))$.

Associated to a tube $T \in \mathcal{T}$, we define the cut-off $\tilde{\chi}_T$ on \mathbb{R}^{n+1} by

$$\tilde{\chi}_T(x, t) = \tilde{\chi}_{D(x_T - t\nabla\varphi(\xi_T), t; c^{-2}r)}(x).$$

We are ready to state the main result of this Section.

Lemma 4.1. *Let Q be a cube of radius $R \gg 1$, let c be such that $R^{-\frac{1}{4}+} \ll c \lesssim 1$ and let $J \in \mathbb{N}$ be such that $r = 2^{-J}R \approx R^{\frac{1}{2}}$. Let $\phi = \mathcal{E}f$ be a free wave with $\text{margin}(\phi) > 0$. For each $T \in \mathcal{T}$ there is a free solution ϕ_T , that is localized in a neighborhood of size $CR^{-\frac{1}{2}}$ of*

the leaf $S(\xi_T)$ and obeying $\text{margin}(\phi_T) \geq \text{margin}(f) - CR^{-\frac{1}{2}}$. The map $f \rightarrow \phi_T$ is linear and

$$(4.2) \quad \phi = \sum_{T \in \mathcal{T}} \phi_T.$$

If $\text{dist}(T, Q) \geq 4R$ then

$$(4.3) \quad \|\phi_T\|_{L^\infty(Q)} \lesssim c^{-C} \text{dist}(T, Q)^{-N} M(\phi)^{\frac{1}{2}}.$$

The following estimates hold true

$$(4.4) \quad \sum_T \sup_{q \in Q_J(Q)} \tilde{\chi}_T(x_q, t_q)^{-N} \|\phi_T\|_{L^2(q)}^2 \lesssim c^{-C} r M(\phi)$$

and

$$(4.5) \quad \left(\sum_{q_0} M \left(\sum_T m_{q_0, T} \phi_T \right) \right)^{\frac{1}{2}} \leq (1 + cC) M(\phi),$$

provided that the coefficients $m_{q_0, T} \geq 0$ satisfy

$$(4.6) \quad \sum_{q_0} m_{q_0, T} = 1, \quad \forall T \in \mathcal{T}.$$

The above construction was provided for the case of conic surfaces (single-conic in our context) in [22], and for more general surfaces in [2]. We sketch the main parts of the argument. We start with the partition

$$\tilde{U}' = \bigcup_{\xi \in \tilde{\mathcal{L}}} A_\xi$$

where A_ξ consists of the points in \tilde{U}' that are closer to ξ than any other elements of $\tilde{\mathcal{L}}$. Therefore A_ξ belongs to the $O(r^{-1})$ neighborhood of ξ .

Let G be the set of all translations in \mathbb{R}^{n-2} by vectors of size at most $O(r^{-1})$; in particular these translations differ from identity by $O(r^{-1})$. Let $d\Omega$ be a smooth compactly supported probability measure on the interior of G . For each $\Omega \in G$ and $\xi_0 \in \mathcal{L}$, we define the Fourier projectors by

$$\mathcal{F}(P_{\Omega, \xi_0} g)(\xi) = \chi_{\Omega(A_{\pi \circ \mathbf{x}}(\xi_0))} (\pi \circ \mathbf{x}(\xi)) \hat{g}(\xi).$$

For fixed $\Omega \in G$, this leads to the decomposition:

$$(4.7) \quad g = \sum_{\xi_0 \in \mathcal{L}} P_{\Omega, \xi_0} g.$$

The terms above have good frequency support: $P_{\Omega, \xi_0} g$ is localized in frequency in a neighborhood of size Cr^{-1} of $U(\xi_0)$; indeed, this follows from the definition of the sets A_ξ , the properties of Ω , the smoothness properties of \mathbf{x} and the fact that $\pi \circ \mathbf{x}(U(\xi_T)) = \pi \circ \mathbf{x}(\xi_T)$.

We note the following: the sole reason for bringing in the additional coordinate map \mathbf{x} into the picture was to accommodate the simple group structure G used above; this is very easy in the context of the fibers $\mathbf{x}(U_\alpha)$, or more precisely their projections $\pi \circ \mathbf{x}(U_\alpha)$ as one simply uses the translations on \mathbb{R}^{n-2} . At the original level of the fibers U_α , this would be more difficult, but probably doable. For instance, for the classical conical model (1.4), one

assigns to each leaf U_α an angle $\omega_\alpha \in \mathbb{S}^{n-2}$ and can use the rotations in $SO(n)$ that leave the \mathbb{S}^{n-2} invariant to construct G . This is essentially the construction used in [22].

Now we proceed with the spatial localization. For each $x_0 \in L$, define

$$\eta^{x_0}(x) = \eta_0\left(\frac{c^2}{r}(x - x_0)\right)$$

and notice that, by the Poisson summation formula and properties of η_0 ,

$$(4.8) \quad \sum_{x_0 \in L} \eta^{x_0} = 1.$$

In Section 3.1 we interpreted $\phi = \mathcal{E}f$ as an evolution with initial data $\phi(0) = \check{f}$. Based on this, we define

$$\phi_T(0) = \eta^{x_T}(x) \int P_{\Omega, \xi_T} \phi(0) d\Omega$$

and evolve this, at all other times, by the free flow

$$\phi_T(t) = e^{it\varphi(D)} \phi_T(0).$$

The statements (4.2)-(4.4) are standard in the wave packet theory, and they rely on the phase-space localization properties of the initial data $\phi_T(0)$ and the properties of the flow $e^{it\varphi(D)}$; the properties of the flow $e^{it\varphi(D)}$ are closely related to the geometry of the characteristic surface S and in our particular case the key property is that in the support of $\hat{\phi}_T(0)$ we have $|N(\Sigma(\xi)) - N(\Sigma(\xi'))| \lesssim r^{-1}$ or, in terms of phase function, $|\nabla\varphi(\xi) - \nabla\varphi(\xi')| \lesssim r^{-1}$. This guarantees that the wave ϕ_T is highly localizes in the tube T . The details can be filled in using similar arguments to those in [2, Lemma 4.1] or [22, Lemma 15.2].

Given that a statement of type (4.5) is less common in a wave packet decomposition, we include a proof of it for convenience. This proof follows closely the one in [22, Lemma 15.2]. By using the unitarity of $e^{it\varphi(D)}$ on L^2 , (4.5) is reduced to the corresponding statement at time $t = 0$. We denote $\phi(0) = g$ and the left hand-side of (4.5) becomes

$$\left(\sum_{q_0} \left\| \int \sum_{\xi_0 \in \mathcal{L}} \sum_{x_0 \in L} m_{q_0, T(x_0, \xi_0)} \eta^{x_0} P_{\Omega, \xi_0} g d\Omega \right\|_{L^2}^2 \right)^{\frac{1}{2}}$$

which is obviously bounded by

$$(4.9) \quad \int \left(\sum_{q_0} \left\| \sum_{\xi_0 \in \mathcal{L}} \sum_{x_0 \in L} m_{q_0, T(x_0, \xi_0)} \eta^{x_0} P_{\Omega, \xi_0} g \right\|_{L^2}^2 \right)^{\frac{1}{2}} d\Omega.$$

We define the set

$$\tilde{Y} := \bigcup_{\xi_0 \in \tilde{\mathcal{L}}} \{ \xi \in A_{\xi_0} : d(\xi, A_{\xi_0}^c) > Cc^2r^{-1} \}$$

and the associated multiplier $P_{\Omega(\tilde{Y})}$ with symbol $\chi_{\Omega(\tilde{Y})} \circ \pi \circ \mathbf{x}(\xi)$, that is

$$\mathcal{F}(P_{\Omega(\tilde{Y})}h) = \chi_{\Omega(\tilde{Y})}(\pi(\mathbf{x}(\xi))) \hat{h}(\xi).$$

Using the triangle inequality we bound the term (4.9) by

$$(4.10) \quad \int \left(\sum_{q_0} \left\| \sum_{\xi_0 \in \mathcal{L}} \sum_{x_0 \in L} m_{q_0, T(x_0, \xi_0)} \eta^{x_0} P_{\Omega, \xi_0} P_{\Omega(\tilde{Y})} g \right\|_{L^2}^2 \right)^{\frac{1}{2}} d\Omega$$

and

$$(4.11) \quad \int \left(\sum_{q_0} \left\| \sum_{\xi_0 \in \mathcal{L}} \sum_{x_0 \in L} m_{q_0, T(x_0, \xi_0)} \eta^{x_0} P_{\Omega, \xi_0} (1 - P_{\Omega(\tilde{Y})}) g \right\|_{L^2}^2 \right)^{\frac{1}{2}} d\Omega$$

Given $\xi_1, \xi_2 \in \mathcal{L}$, $P_{\Omega, \xi_1} P_{\Omega(\tilde{Y})} g$ and $P_{\Omega, \xi_2} P_{\Omega(\tilde{Y})} g$ have Fourier support in the sets $\mathbf{x}^{-1}(\Omega(A_{\pi \mathbf{x}(\xi_1)} \setminus \tilde{Y}))$ and $\mathbf{x}^{-1}(\Omega(A_{\pi \mathbf{x}(\xi_2)} \setminus \tilde{Y}))$, respectively. The two sets $\Omega(A_{\pi \mathbf{x}(\xi_1)} \setminus \tilde{Y})$ and $\Omega(A_{\pi \mathbf{x}(\xi_2)} \setminus \tilde{Y})$ are at distance $\geq Cc^2r^{-1}$ and the map \mathbf{x}^{-1} may change this distance by a factor, that is the distance between the sets $\mathbf{x}^{-1}(\Omega(A_{\pi \mathbf{x}(\xi_1)} \setminus \tilde{Y}))$ and $\mathbf{x}^{-1}(\Omega(A_{\pi \mathbf{x}(\xi_2)} \setminus \tilde{Y}))$ is $\geq Cc^2r^{-1}$, after redefining C .

The functions η^{x_0} have Fourier support in the set $|\xi| \leq c^2r^{-1}$, thus the Fourier supports of $\eta^{x_1} P_{\Omega, \xi_1} P_{\Omega(\tilde{Y})} g$ and $\eta^{x_2} P_{\Omega, \xi_2} P_{\Omega(\tilde{Y})} g$ are also at distance $\geq Cc^2r^{-1}$; recall that $\mathcal{F}(\eta^{x_1} P_{\Omega, \xi_1} P_{\Omega(\tilde{Y})} g) = \hat{\eta}^{x_1} * \mathcal{F}(P_{\Omega, \xi_1} P_{\Omega(\tilde{Y})} g)$. The conclusion to draw from this is that if $\xi_1 \neq \xi_2$, then $\eta^{x_1} P_{\Omega, \xi_1} P_{\Omega(\tilde{Y})} g$ and $\eta^{x_2} P_{\Omega, \xi_2} P_{\Omega(\tilde{Y})} g$ have disjoint Fourier support. As a consequence, the term in (4.10) equals

$$\begin{aligned} &= \int \left(\sum_{q_0} \sum_{\xi_0 \in \mathcal{L}} \left\| \sum_{x_0 \in L} m_{q_0, T(x_0, \xi_0)} \eta^{x_0} P_{\Omega, \xi_0} P_{\Omega(\tilde{Y})} g \right\|_{L^2}^2 \right)^{\frac{1}{2}} d\Omega \\ &= \int \left(\sum_{\xi_0 \in \mathcal{L}} \int |P_{\Omega, \xi_0} P_{\Omega(\tilde{Y})} g(x)|^2 \sum_{q_0} \left(\sum_{x_0 \in L} m_{q_0, T(x_0, \xi_0)} \eta^{x_0}(x) \right)^2 dx \right)^{\frac{1}{2}} d\Omega \\ &\leq \int \left(\sum_{\xi_0 \in \mathcal{L}} \int |P_{\Omega, \xi_0} P_{\Omega(\tilde{Y})} g(x)|^2 \left(\sum_{q_0} \sum_{x_0 \in L} m_{q_0, T(x_0, \xi_0)} \eta^{x_0}(x) \right)^2 dx \right)^{\frac{1}{2}} d\Omega \end{aligned}$$

Using (4.6) and (4.8) we can continue with the

$$= \int \left(\sum_{\xi_0 \in \mathcal{L}} \int |P_{\Omega, \xi_0} P_{\Omega(\tilde{Y})} g(x)|^2 dx \right)^{\frac{1}{2}} d\Omega \leq \int \|P_{\Omega(\tilde{Y})} g\|_{L^2} d\Omega \leq \|g\|_{L^2},$$

where in passing to the second inequality we have used the orthogonality of the projectors P_{Ω, ξ_0} with respect to ξ_0 .

We use a similar argument for dealing with the term in (4.11), except that now we invoke only almost orthogonality arguments, and not full orthogonality; therefore we can bound the term in (4.11) by

$$\lesssim \int \|(1 - P_{\Omega(\tilde{Y})}) g\|_{L^2} d\Omega.$$

We use the Cauchy-Schwartz inequality and Plancherel to continue the sequence of bounds,

$$\begin{aligned}
&\lesssim \left(\int \| (1 - P_{\Omega(\tilde{Y})})g \|_{L^2}^2 d\Omega \right)^{\frac{1}{2}} \\
&= \left(\int \| (1 - \chi_{\Omega(\tilde{Y})}(\pi \mathbf{x}(\xi))) \hat{g}(\xi) \|_{L^2}^2 d\Omega \right)^{\frac{1}{2}} \\
&= \left(\int |\hat{g}(\xi)|^2 \int (\chi_{\Omega(\tilde{Y})^c}(\pi \mathbf{x}(\xi)))^2 d\Omega d\xi \right)^{\frac{1}{2}}
\end{aligned}$$

For fixed ξ , the integral with respect to Ω is $\lesssim c^2$, thus the final bound is

$$\lesssim c \|g\|_{L^2}.$$

Combining the two bounds we derived for (4.10) and (4.11) leads to (4.5).

5. LOCALIZATION OF THE MULTILINEAR ESTIMATE

In [1] we established a refinement of the generic trilinear restriction estimate under the additional assumption of small support of one of the terms involved, see (5.3) below. In this section we provide a slight modification of this result which is needed for technical reasons.

We recall the setup from [1]. We are given 3 smooth hypersurfaces $S_i = \Sigma_i(U_i)$ with smooth parameterizations Σ_i . We assume that there exists $\nu > 0$ such that

$$(5.1) \quad \text{vol}(N_1(\zeta_1), N_2(\zeta_2), N_3(\zeta_3)) \geq \nu,$$

for all choices $\zeta_i \in \Sigma_i(U_i)$. Here by $\text{vol}(N_1(\zeta_1), N_2(\zeta_2), N_3(\zeta_3))$ we mean the volume of the 3-dimensional parallelepiped spanned by the vectors $N_1(\zeta_1), N_2(\zeta_2), N_3(\zeta_3)$.

Assume that $\Sigma_1(\text{supp} f_1) \subset B(\mathcal{H}, \mu)$, where $B(\mathcal{H}, \mu)$ is the neighborhood of size μ of the 3-dimensional affine subspace \mathcal{H} . Assume that $|N_1(\zeta_1) - \pi_{\mathcal{H}} N_1(\zeta_1)| \lesssim \mu, \forall \zeta_1 \in \Sigma_1(\text{supp} f_1)$, where $\pi_{\mathcal{H}} : \mathbb{R}^{n+1} \rightarrow \mathcal{H}$ is the projection onto \mathcal{H} . In addition assume that if $N_i, i = 4, \dots, n+1$ is a basis of the normal space \mathcal{H}^\perp to \mathcal{H} , then $N_1(\zeta_1), \dots, N_3(\zeta_3), N_4, \dots, N_{n+1}$ are transversal in the following sense:

$$(5.2) \quad |\det(N_1(\zeta_1), N_2(\zeta_2), N_3(\zeta_3), N_4, \dots, N_{n+1})| \geq \nu,$$

for all choices $\zeta_i \in \Sigma_i(U_i)$.

Under these hypothesis we proved that

$$(5.3) \quad \|\mathcal{E}_1 f_1 \mathcal{E}_2 f_2 \mathcal{E}_3 f_3\|_{L^1(B(0,r))} \leq C(\epsilon) \mu^{\frac{n-2}{2}} r^\epsilon \|f_1\|_{L^2(U_1)} \|f_2\|_{L^2(U_2)} \|f_3\|_{L^2(U_3)}.$$

In this section we show that this estimate localizes in the following sense: assume q is a cube of size r and that $\mu \gtrsim r^{-1}$ then

$$(5.4) \quad \|\mathcal{E}_1 f_1 \mathcal{E}_2 f_2 \mathcal{E}_3 f_3\|_{L^1(q)} \leq C(\epsilon) r^\epsilon \mu^{\frac{n-2}{2}} r^{-\frac{3}{2}} \prod_{i=1}^3 \|\tilde{\chi}_q \mathcal{E}_i f_i\|_{L^2}$$

There are few aspects in the above estimate that need to be highlighted. Switching from balls of radius r to cubes of radius r makes no difference. The estimate says that inside q the main contributions come from $\mathcal{E}_i f_i$ inside a dilate of q . This is simply a consequence of the finite speed of propagation. The factor $r^{-\frac{3}{2}}$ exhibits an apparent improvement over (5.3). This is explained by the fact that while in (5.3), the norms of $\mathcal{E}_i f_i$ are measured along hyperplanes, in (5.4) the norms of $\mathcal{E}_i f_i$ are measured on cubes of size r and the additional dimension explains

the additional factors. Finally, the condition $\mu \gtrsim r^{-1}$ is crucial in using the localization on the physical side, without altering the localization at scale $\lesssim \mu$ of $\Sigma_1(\text{supp} f_1)$.

To prove (5.4), we need to invoke the localization machinery developed in [1]. For each $i = 1, 2, 3$ we fix $\zeta_i^0 \in \Sigma_i(\text{supp} f_i)$, $N_i = N_i(\zeta_i^0)$ and let \mathcal{H}_i be the hyperplane on the physical side passing through the origin with normal N_i . We denote by π_{N_i} the projection onto \mathcal{H}_i along N_i . Then, we choose a basis $N_i, i = 4, \dots, n+1$ of the normal plane to \mathcal{H} , let \mathcal{H}_i be the hyperplane on the physical side passing through the origin with normal N_i and denote by π_{N_i} the projection onto \mathcal{H}_i along N_i . The set $\{N_i\}_{i=1, \dots, n+1}$ is a basis of \mathbb{R}^{n+1} .

In each $\mathcal{H}_i, i = 1, 2, 3$, we construct the set $\mathcal{CH}_i(r)$ to be the set of n -dimensional cubes of size r . For $y_i \in \mathbb{R}$, we define $\mathcal{H}_i + y_i N_i$ to be the translate of \mathcal{H}_i by $y_i N_i$ (in the normal direction). Then $\mathcal{CH}_i(r) + y_i N_i$ is the corresponding translate of $\mathcal{CH}_i(r)$ by $y_i N_i$.

For $i \in \{1, 2, 3\}$, $r > 0$ and for a cube $q \in \mathcal{H}_i$ of radius r , we define $\chi_q : \mathcal{H}_i \rightarrow \mathbb{R}$ by

$$\chi_q(x) = \eta_0\left(\frac{x - c(q)}{r}\right)$$

Notice that $\mathcal{F}_i \chi_q$ has Fourier support in the ball of radius $\lesssim r^{-1}$. This object is very similar to the ones defined in the Notation Section 1.1, the only difference is that they localize cubes living in different linear spaces. By the Poisson summation formula and properties of η_0 ,

$$(5.5) \quad \sum_{q \in \mathcal{CH}_i(r)} \chi_q = 1.$$

Using the properties of χ_q , a direct exercise shows that for each $N \in \mathbb{N}$, the following holds true

$$(5.6) \quad \sum_{q \in \mathcal{CH}_i(r)} \left\| \left\langle \frac{x - c(q)}{r} \right\rangle^N \chi_q g \right\|_{L^2}^2 \lesssim_N \|g\|_{L^2}^2$$

for any $g \in L^2(\mathcal{H}_i)$. Here, the variable x is the argument of g and belongs to \mathcal{H}_i .

Now we can start the argument of for (5.4). Given any vector $y \in \mathbb{R}^{n+1}$ with $|y_i - c_i(q)| \leq r, i = 1, 2, 3$ and $y_i = c_i(q), 4 \leq i \leq n+1$, we claim the following:

$$(5.7) \quad \begin{aligned} & \|\Pi_{i=1}^3 \mathcal{E}_i f_i\|_{L^1(q)} \lesssim C(\epsilon) r^\epsilon \mu^{\frac{n-2}{2}} \\ & \cdot \Pi_{i=1}^3 \left(\sum_{q' \in \mathcal{CH}_i(r) + y_i N_i} \left\langle \frac{d(\pi_{N_i} q, q')}{r} \right\rangle^{-N} \left\| \left\langle \frac{x - c(q')}{r} \right\rangle^N \chi_{q'} \mathcal{E}_i f_i \right\|_{L^2(\mathcal{H}_i + y_i N_i)}^2 \right)^{\frac{1}{2}} \end{aligned}$$

This allows us to average over the values of (y_1, y_2, y_3) satisfying $|y_i - c_i(q)| \leq r$ (keeping $y_i = c_i(q), i \geq 4$) to further bound

$$\begin{aligned} & \|\Pi_{i=1}^3 \mathcal{E}_i f_i\|_{L^1(q)} \lesssim C(\epsilon) r^\epsilon \mu^{\frac{n-2}{2}} r^{-\frac{3}{2}} \\ & \cdot \Pi_{i=1}^3 \left(\int_{|y_i| \leq r} \sum_{q' \in \mathcal{CH}_i(r) + y_i N_i} \left\langle \frac{d(\pi_{N_i} q, q')}{r} \right\rangle^{-N} \left\| \left\langle \frac{x - c(q')}{r} \right\rangle^N \chi_{q'} \mathcal{E}_i f_i \right\|_{L^2(\mathcal{H}_i + y_i N_i)}^2 \right)^{\frac{1}{2}} \\ & \lesssim C(\epsilon) r^\epsilon \mu^{\frac{n-2}{2}} r^{-\frac{3}{2}} \Pi_{i=1}^3 \|\tilde{\chi}_q \mathcal{E}_i f_i\|_{L^2} \end{aligned}$$

This gives the estimate (5.4). All that is left is the justification of (5.7). We note that this type of a posteriori improvement of the multilinear estimate was at the core of the arguments in [1] and that, while (5.7) is not explicitly established there, its proof is essentially derived along the same lines as (3.6) in that paper.

We first note that, without restricting the generality of the argument we can assume that $c(q) = 0$. Following the argument below, it will be clear that the proof for $y = 0$ extends to the more general case when $|y_i| \leq r, i = 1, 2, 3$.

We fix $i = 1$. In [1] we explained that the problem is reducible to the case when $\Sigma_1 : U_1 \subset \mathcal{H}_1 \rightarrow \mathbb{R}^{n+1}$ with $\Sigma_1(\xi') = (\xi', \varphi_1(\xi'))$ and $\text{supp} f_1 \subset B(\mathcal{H} \cap \mathcal{H}_1, \mu)$, the μ neighborhood of $\mathcal{H} \cap \mathcal{H}_1 \subset \mathcal{H}_1$. Then we have the representation

$$(5.8) \quad \mathcal{E}_1 f(x_1, x') = \int_{U_1} e^{i(x' \xi' + x_1 \varphi_1(\xi'))} f(\xi') d\xi'.$$

We highlight a commutator estimate which is needed due to the uncertainty principle. For any fixed $x_0 \in \mathbb{R}^{n+1}$, it holds true that

$$(5.9) \quad (x' - x'_0 - x_1 \nabla \varphi_1(\frac{D'}{i}))^N \mathcal{E}_1 f_1 = \mathcal{E}_1(\mathcal{F}_1((x' - x'_0)^N \mathcal{F}_1^{-1} f_1)), \quad \forall N \in \mathbb{N}.$$

This is a direct computation using (5.8) and it suffices to check it for $N = 1$.

Fix $q' \in \mathcal{CH}_1(r) + y$. With the notation $A = C(\epsilon) \mu^{\frac{n-2}{2}} r^\epsilon$ we have

$$\begin{aligned} & \| (x' - c(q') - x_1 \nabla \varphi_1(\xi'_0)) \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^3 \mathcal{E}_i f_i \|_{L^1(q)} \\ &= \| (x' - c(q') - x_1 \nabla \varphi_1(\xi')) \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^3 \mathcal{E}_i f_i \|_{L^1(q)} \\ &+ \| x_1 (\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\xi')) \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^3 \mathcal{E}_i f_i \|_{L^1(q)} \\ &= \| \mathcal{E}_1 \mathcal{F}_1((x' - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^3 \mathcal{E}_i f_i \|_{L^1(q)} \\ &+ \| x_1 \mathcal{E}_1 \mathcal{F}_1((\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\xi')) \chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^3 \mathcal{E}_i f_i \|_{L^1(q)} \\ &\leq A (\| (x' - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2} + r \| (\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\xi')) \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2}) \Pi_{i=2}^3 \| f_i \|_{L^2} \\ &\lesssim A (\| (x' - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2} + r \| \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2}) \Pi_{i=2}^3 \| f_i \|_{L^2} \\ &\lesssim r A \| \langle \frac{x' - c(q')}{r} \rangle \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2} \Pi_{i=2}^3 \| f_i \|_{L^2} \end{aligned}$$

We have used the following: (5.9) in justifying the equality between the terms on the second and fourth line, the induction hypothesis and the fact that inside Q we have $|x_1| \lesssim r$ to justify the inequality in the sixth line (note that this part extends easily for more general y_1 provided $|y_1| \leq r$). Since $\chi_{q'}$ has Fourier support in a ball of radius $\lesssim r^{-1} \leq \mu$, $\mathcal{E}_1 \mathcal{F}_1((x' - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1)$ keeps the localization property of f_1 , that is $\text{supp} \mathcal{F}_1((x' - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1) \subset B(\mathcal{H} \cap \mathcal{H}_1, C\mu)$ for some fixed constant C .

For $x' \in \pi_{N_1}(q)$, it holds that $\langle \frac{x' - c(q') - x_1 \nabla \varphi_1(\xi'_0)}{r} \rangle \approx \langle \frac{d(\pi_{N_1}(q), q')}{r} \rangle$. This is justified by the fact that $|x_1| \lesssim r$ and $|\nabla \varphi_1(\xi'_0)| \leq 1$, therefore the contribution of $|x_1 \nabla \varphi_1(\xi'_0)| \lesssim r$ is negligible. From this and the previous set of estimates, we conclude that

$$\| \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^3 \mathcal{E}_i f_i \|_{L^1(q)} \lesssim A \langle \frac{d(\pi_{N_1}(q), q')}{r} \rangle^{-1} \| \langle \frac{x' - c(q')}{r} \rangle \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2} \Pi_{i=2}^3 \| f_i \|_{L^2}$$

Repeating the argument gives

$$\|\mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^3 \mathcal{E}_i f_i\|_{L^1(q)} \lesssim_N A \langle \frac{d(\pi_{N_1} q, q')}{r} \rangle^{-N} \|\langle \frac{x' - c(q')}{r} \rangle^N \chi_{q'} \mathcal{F}_1^{-1} f_1\|_{L^2} \Pi_{i=2}^3 \|f_i\|_{L^2}$$

Using (5.5) and the above, we obtain

$$\begin{aligned} & \|\mathcal{E}_1 f_1 \cdot \Pi_{i=2}^{n+1} \mathcal{E}_i f_i\|_{L^1(q)} \\ & \leq A \sum_{q' \in \mathcal{CH}_1(r)} \|\mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^3 \mathcal{E}_i f_i\|_{L^1(q)} \\ & \lesssim_N A \left(\sum_{q' \in \mathcal{CH}_1(r)} \langle \frac{d(\pi_{N_1} q, q')}{r} \rangle^{-N} \|\langle \frac{x' - c(q')}{r} \rangle^N \chi_{q'} \mathcal{F}_1^{-1} f_1\|_{L^2} \right) \Pi_{i=2}^3 \|f_i\|_{L^2} \\ & \lesssim_N A \left(\sum_{q' \in \mathcal{CH}_1(r)} \langle \frac{d(\pi_{N_1} q, q')}{r} \rangle^{-(2N-4)} \|\langle \frac{x' - c(q')}{r} \rangle^N \chi_{q'} \mathcal{F}_1^{-1} f_1\|_{L^2}^2 \right)^{\frac{1}{2}} \Pi_{i=2}^3 \|f_i\|_{L^2}. \end{aligned}$$

This is the improvement we claimed in (5.7) for $\mathcal{E}_1 f_1$. The improvements for the $\mathcal{E}_2 f_2$ and $\mathcal{E}_3 f_3$ are obtained in a similar manner.

6. TABLE CONSTRUCTION AND THE INDUCTION ARGUMENT

This section contains the main argument for the proof of Theorem 1.1. In Proposition 6.1 we construct tables on cubes: this is a way of re-organizing the information on one term, say ϕ_1 , at smaller scales based on information from one of the other interacting terms, ϕ_2 or ϕ_3 . This type of argument is inspired by the work on the conic surfaces of Tao in [22]. Based on this table construction, we will prove the inductive bound claimed in Proposition 3.2.

Proposition 6.1. *Let Q be a cube of size $R \gg 2^{2C_0}$. Assume $\phi_i = \mathcal{E}_i f_i, i \in \{1, 2, 3\}$ have positive margin. Then there is a table $\Phi_1 = \Phi_c(\phi_1, \phi_2, Q)$ with depth C_0 such that the following properties hold true:*

$$(6.1) \quad \phi_1 = \sum_{q \in \mathcal{Q}_{C_0}(Q)} \Phi_1^{(q)},$$

$$(6.2) \quad \text{margin}(\Phi) \geq \text{margin}(\phi) - CR^{-\frac{1}{2}}.$$

$$(6.3) \quad M(\Phi) \leq (1 + cC)M(\phi),$$

and for any $q', q'' \in \mathcal{Q}_{C_0}(Q), q' \neq q''$

$$(6.4) \quad \|\Phi_1^{(q')} \phi_2 \phi_3\|_{L^1((1-c)q'')} \lesssim_\epsilon c^{-C} R^{-\frac{n-2}{4} + \epsilon} \Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i).$$

Remark 1. *The above result is stated for scalar ϕ_1, ϕ_2, ϕ_3 , but it holds for vector versions as well. Most important is that we can construct $\Phi_1 = \Phi_c(\phi_1, \Phi_2, Q)$ where Φ_2 is a vector free wave and all its scalar components satisfy similar properties to the ϕ_2 above.*

Remark 2. *We note that $\Phi_1 = \Phi_c(\phi_1, \phi_2, Q)$ means that the table Φ_1 is constructed from ϕ_1 , which is natural in light of (6.1), and ϕ_2 . But it does not depend on ϕ_3 . Obviously, we could have constructed it from ϕ_1 and ϕ_3 , ending with a different object.*

Proof. There are several scales involved in this argument. The large scale is the size R of the cube Q . The coarse scale is $2^{-C_0}R \gg R^{\frac{1}{2}}$, this being the size of the smaller cubes in $\mathcal{Q}_{C_0}(Q)$ and the subject of the claims in the Proposition. Then there is the fine scale $r = 2^{-j}R$ chosen such that $r \approx R^{\frac{1}{2}}$. Notice that r is the proper scale for wave packets corresponding to time scales R and also that their scale is $c^{-2}r \ll 2^{-C_0}R$, last one being the scale of cubes in $\mathcal{Q}_{C_0}(Q)$.

We use Lemma 4.1 with $J = j$ to construct the wave packet decomposition for ϕ_1 ,

$$\phi_1 = \sum_{T_1 \in \mathcal{T}_1} \phi_{1,T_1}.$$

For any $q_0 \in \mathcal{Q}_{C_0}(Q)$ and $T_1 \in \mathcal{T}_1$ we define

$$m_{q_0, T_1} := \|\tilde{\chi}_{T_1} \phi_2\|_{L^2(q_0)}^2$$

and

$$m_{T_1} := \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} m_{q_0, T_1}.$$

Based on this we define

$$(6.5) \quad \Phi_1^{(q_0)} := \sum_{T_1 \in \mathcal{T}_1} \frac{m_{q_0, T_1}}{m_{T_1}} \phi_{1, T_1}.$$

By combing the definitions above with the decomposition property (4.2), we obtain

$$\phi_1 = \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} \Phi_1^{(q_0)}$$

thus justifying (6.1).

The margin estimate (6.2) follows from the margin estimate on tubes provided by Lemma 4.1. The coefficients $m_{q_0, T}$ satisfy (4.6), thus the estimate (6.3) follows from (4.5).

All that is left to prove is (6.4), which is equivalent to

$$(6.6) \quad \sum_{q \in \mathcal{Q}_j(Q): d(q, q_0) \gtrsim cR} \|\Phi_1^{(q_0)} \phi_2 \phi_3\|_{L^1(q)} \lesssim_\epsilon c^{-C} r^{-\frac{n-2}{2} + \epsilon} \Pi_{i=1}^3 M(\phi_i).$$

Note that the cubes q are selected at the finer scale dictated the size of cubes in $\mathcal{Q}_j(Q)$. From the definition of $\Phi^{(q_0)}$ in (6.5) we can discard the tubes q which do not intersect $4Q$ based on (4.3), in the sense that their contribution to (6.6) will give a better estimate.

For the tubes intersecting $4Q$, we make another simplification motivated by (4.4) and focus on the tubes which intersect q , that is we focus on the following term

$$\sum_{q \in \mathcal{Q}_j(Q): d(q, q_0) \gtrsim cR} \left\| \sum_{T_1 \cap q \neq \emptyset} \frac{m_{q_0, T_1}}{m_{T_1}} \phi_{1, T_1} \phi_2 \phi_3 \right\|_{L^1(q)}$$

Essentially (4.4) says that the other tubes have off-diagonal type contribution, that is there are enough gains in the case $T_1 \cap q = \emptyset$ to perform any summation. Note that in order to keep the notation simple, we skipped writing $T_1 \in \mathcal{T}_1$ and will do so throughout the rest of this proof.

Using the localized form of the trilinear estimate (5.4), we obtain

$$\|\Phi_1^{(q_0)} \phi_2 \phi_3\|_{L^1(q)} \lesssim_\epsilon r^{-\frac{n+1}{2} + \epsilon} \sum_{T_1} \frac{m_{q_0, T_1}}{m_{T_1}} \|\phi_{1, T_1} \tilde{\chi}_q\|_{L^2} \|\phi_2 \tilde{\chi}_q\|_{L^2} \|\phi_3 \tilde{\chi}_q\|_{L^2}.$$

Here we used that ϕ_{1,T_1} is a free wave of type $\mathcal{E}_1 g_1$ (for some g_1) that is localized in frequency in the $\mu = r^{-1}$ neighborhood of the 2-dimensional affine plane containing the leaf $S_{1,\alpha}$ where we choose the leaf such that $\Sigma_1(\xi_T) \in S_{1,\alpha}$. If we add to the leaf the normal directions $N_1(\Sigma(\xi_T))$, we obtain a 3-dimensional affine subspace \mathcal{H} with the properties required to invoke (5.4).

Using the obvious inequality $\frac{m_{q_0,T_1}}{m_{T_1}} \leq \frac{m_{q_0,T_1}^{\frac{1}{2}}}{m_{T_1}^{\frac{1}{2}}}$, we obtain:

$$\begin{aligned} & \sum_{T_1 \cap q \neq \emptyset} \frac{m_{q_0,T_1}}{m_{T_1}} \|\phi_{1,T_1} \tilde{\chi}_q\|_{L^2} \\ & \lesssim \left(\sum_{T_1} \frac{\|\phi_{1,T_1} \tilde{\chi}_q\|_{L^2}^2}{m_{T_1} \tilde{\chi}_{T_1}(x_q, t_q)} \right)^{\frac{1}{2}} \left(\sum_{T_1} m_{q_0,T} \tilde{\chi}_{T_1}(x_q, t_q) \right)^{\frac{1}{2}}. \end{aligned}$$

Next we claim the following estimate

$$(6.7) \quad \sum_{T_1 \in \mathcal{T}_1} m_{q_0,T_1} \tilde{\chi}_{T_1}(x_q, t_q) \lesssim \|\tilde{\chi}_{S(q)} \phi_2\|_{L^2}^2.$$

Using the definition of m_{q_0,T_1} we identify the function

$$\tilde{\chi}_{S(q)} = \left(\sum_{T_1 \in \mathcal{T}_1} \tilde{\chi}(x_q, t_q) \tilde{\chi}_{T_1} \right) \chi_{q_0}$$

which makes (6.7) hold true. Then we note that $\tilde{\chi}_{S(q)}$ has the following decay property:

$$\tilde{\chi}_{S(q)}(x, t) \lesssim c^{-4} \left(1 + \frac{d((x, t), S(q))}{c^{-2}r} \right)^{-N}.$$

Here the surface $S(q)$ is the translate by $c(q)$ of the neighborhood of size r of cone of normals at S_1 , which we denote by $\mathcal{CN}_1 := \{\alpha N_1(\zeta), \zeta \in S_1, \alpha \in \mathbb{R}\}$. It is important to note that we do not consider the whole cone but only the part with $cR \leq \alpha \lesssim R$. Then the estimate is a consequence of the fact that the tubes T_1 passing thorough q separate inside q_0 and of the separation between q and q_0 , which is quantified by $d(q, q_0) \gtrsim cR$. Quantitatively speaking, given a point in q_0 close to $S(q)$, there are $\lesssim c^{-4}$ tubes T_1 passing through the point and q - this follows from the dispersion estimate (4.1) and the geometry of the family of tubes \mathcal{T}_1 .

Next we claim the following estimate:

$$(6.8) \quad \sum_q \|\tilde{\chi}_{S(q)} \phi_2\|_{L^2} \left(\sum_{T_1 \cap q \neq \emptyset} \frac{\|\phi_{1,T_1} \tilde{\chi}_q\|_{L^2}^2}{m_{T_1} \tilde{\chi}_{T_1}(x_q, t_q)} \right)^{\frac{1}{2}} \|\phi_2 \tilde{\chi}_q\|_{L^2} \|\phi_3 \tilde{\chi}_q\|_{L^2} \lesssim c^{-C} r^{\frac{3}{2}} \Pi_{i=1}^3 M(\phi_i).$$

Combing (6.8) with (6.7) gives (6.6) and this concludes the proofs of all claims of the Proposition.

All that is left to prove is (6.8) and this can be broken down into the following two claims:

$$(6.9) \quad \sum_q \|\phi_2 \tilde{\chi}_{S(q)}\|_{L^2}^2 \|\phi_3 \tilde{\chi}_q\|_{L^2}^2 \lesssim c^{-C} r^2 M(\phi_2) M(\phi_3)$$

and

$$(6.10) \quad \sum_q \left(\sum_{T_1 \cap q \neq \emptyset} \frac{\|\phi_{1,T_1} \tilde{\chi}_q\|_{L^2}^2}{m_{T_1} \tilde{\chi}_{T_1}(x_q, t_q)} \right) \|\phi_2 \tilde{\chi}_q\|_{L^2}^2 \lesssim c^{-C} r M(\phi_1).$$

The proof of (6.10) is similar to the one we used in the bilinear theory, see the proof of the corresponding theorem [2]. By rearranging the sum, it suffices to show

$$\sum_{T_1} \sum_{q \cap T_1 \neq \emptyset} \frac{\|\phi_{1,T_1} \tilde{\chi}_q\|_{L^2}^2 \|\phi_2 \tilde{\chi}_q\|_{L^2}^2}{m_{T_1} \tilde{\chi}_{T_1}(x_q, t_q)} \lesssim r M(\phi_1).$$

The inner sum is estimated as follows:

$$\sum_{q \cap T_1 \neq \emptyset} \frac{\|\phi_2 \tilde{\chi}_q\|_{L^2}^2}{m_{T_1} \tilde{\chi}_{T_1}(x_q, t_q)} \lesssim \frac{\|\phi_2 \tilde{\chi}_{T_1}\|_{L^2}^2}{m_{T_1}} \lesssim 1,$$

and the outer one is estimated by

$$\sum_{T_1} \sup_q \|\phi_{1,T_1} \tilde{\chi}_q\|_{L^2}^2 \lesssim r \sum_{T_1} M(\phi_{1,T_1}) \lesssim r M(\phi_1),$$

which is obvious given the size of q in the x_1 -direction is $\approx r$ and the mass of ϕ_{1,T_1} is constant across slices in space with $x_1 = \text{constant}$.

We continue with the proof of (6.9). Using the fast decay of $\tilde{\chi}_{S(q)}$ away from $S(q)$ and of $\tilde{\chi}_q$ away from q , it suffices to show that

$$(6.11) \quad \sum_q \|\chi_{S(q)}^c \phi_2\|_{L^2}^2 \|\chi_q^c \phi_3\|_{L^2}^2 \lesssim r^2 M(\phi_2) M(\phi_3),$$

where by $\chi_{S(q)}^c$ we denote the characteristic function of the set $S(q)$ and recall that χ_q^c is the characteristic function of q ; this will come at the cost of picking a factor of c^{-C} .

We define the following relation: $q' \sim q$ if $q' \cap S(q) \neq \emptyset$ and note that this is equivalent to saying that there is a tube $T_1 \in \mathcal{T}_1$ intersecting both q and q' and that $d(q, q') \gtrsim cR$. We start from the obvious inequality

$$\|\chi_{S(q)}^c \phi_2\|_{L^2}^2 \lesssim \sum_{q' \sim q} \|\chi_{q'}^c \phi_2\|_{L^2}^2$$

which implies

$$\sum_q \|\chi_{S(q)}^c \phi_2\|_{L^2}^2 \|\chi_q^c \phi_3\|_{L^2}^2 \lesssim \sum_q \sum_{q' \sim q} \|\chi_{q'}^c \phi_2\|_{L^2}^2 \|\chi_q^c \phi_3\|_{L^2}^2$$

We are tacitly using again at this point the full dispersion property of the set of normals \mathcal{N}_1 : the tubes T_1 passing through q separate inside q_0 ; in the absence of this property, the above inequality would fail, as we would encounter large number of tubes $T_1 \in \mathcal{T}_1$ passing through both q and q' .

At this point we use a wave packet decomposition for ϕ_2 and ϕ_3 : we invoke again Proposition 4.1, but this time with $c \approx 1$. Using (4.4), we reduce (6.11) to the following

$$(6.12) \quad \sum_q \sum_{q' \sim q} \left(\sum_{T_2 \cap q' \neq \emptyset} M(\phi_{T_2}) \right) \left(\sum_{T_3 \cap q \neq \emptyset} M(\phi_{T_3}) \right) \lesssim M(\phi_2) M(\phi_3).$$

Here we skip writing down that the sum run over $T_2 \in \mathcal{T}_2$ and $T_3 \in \mathcal{T}_3$ for keeping the notation shorter and we will do so for the rest of the argument.

The key point in justifying (6.12) is that, as we vary (q, q') with $q' \sim q$, the number of occurrences of a pair of tubes (T_2, T_3) on the left hand-side is bounded by a universal constant. Indeed, if that is the case we use (4.5) to obtain the bound

$$\sum_q \sum_{q' \sim q} \left(\sum_{T_2 \cap q' \neq \emptyset} M(\phi_{T_2}) \right) \left(\sum_{T_3 \cap q \neq \emptyset} M(\phi_{T_3}) \right) \lesssim \sum_{T_2 \in \mathcal{T}_2} \sum_{T_3 \in \mathcal{T}_3} M(\phi_{T_2}) M(\phi_{T_3}) \lesssim M(\phi_2) M(\phi_3).$$

We finish the argument by establishing an upper bound on the number of occurrences of a pair of tubes (T_2, T_3) on the left-hand side of (6.12). Assume that a pair (T_2, T_3) shows up multiple times. That means that there are $(q, q'), (\tilde{q}, \tilde{q}')$ such that $q \sim q', \tilde{q} \sim \tilde{q}'$ and $q' \cap T_2 \neq \emptyset, \tilde{q}' \cap T_2 \neq \emptyset, q \cap T_3 \neq \emptyset, \tilde{q} \cap T_3 \neq \emptyset$. We tolerate repeated occurrences coming from the setup $d(q, \tilde{q}) \lesssim r$ and $d(q', \tilde{q}') \lesssim r$, which are bounded by a universal constant, but rule out all the others.

Consider the case $d(q, \tilde{q}), d(q', \tilde{q}') \gg r$. This implies the following:

$$c(q) - c(\tilde{q}) = \alpha_3 N_3 + O(r) \text{ for some } N_3 = N_3(\zeta_3), \zeta_3 \in S_3, |\alpha_3| \gg r,$$

$$c(q') - c(\tilde{q}') = \alpha_2 N_2 + O(r) \text{ for some } N_2 = N_2(\zeta_2), \zeta_2 \in S_2, |\alpha_2| \gg r,$$

$$c(q) - c(q') = \alpha_1 N_1 + O(r), c(\tilde{q}) - c(\tilde{q}') = \tilde{\alpha}_1 \tilde{N}_1 + O(r), N_1, \tilde{N}_1 \in \mathcal{CN} \text{ } |\alpha_1|, |\tilde{\alpha}_1| \gg r.$$

Since $c(q) - c(q') - (c(\tilde{q}) - c(\tilde{q}')) = c(q) - c(\tilde{q}) - (c(q') - c(\tilde{q}'))$, this implies

$$\alpha_1 N_1 - \tilde{\alpha}_1 \tilde{N}_1 = \alpha_3 N_3 - \alpha_2 N_2 + O(r)$$

Since $\alpha_1 N_1 - \tilde{\alpha}_1 \tilde{N}_1 \in d\text{span} \mathcal{N}_1$ we can invoke the result in Lemma 2.1 to obtain

$$|\alpha_1 N_1 - \tilde{\alpha}_1 \tilde{N}_1 - \alpha_3 N_3 + \alpha_2 N_2| \geq \max(|\alpha_2|, |\alpha_3|) \gg r$$

which is in contradiction with the previous statement.

The other two cases, $d(q, \tilde{q}) \gg r, d(q', \tilde{q}') \lesssim r$ and $d(q, \tilde{q}) \lesssim r, d(q', \tilde{q}') \gg r$ are ruled out in a similar way: one notices that in the above proof we needed only one of coefficients α_2, α_3 to have large absolute value. □

Proof of Proposition 3.2. Let $\phi_i = \mathcal{E}_i f_i$ satisfying the margin requirements (3.2). Let Q_R be an arbitrary cube of radius R . From Lemma 3.3 it follows that there is a cube $Q \subset 4Q_R$ of size $2R$ such that

$$(6.13) \quad \|\phi_1 \phi_2 \phi_3\|_{L^p(Q_R)} \leq (1 + cC) \|\phi_1 \phi_2 \phi_3\|_{L^p(I^{c,j}(Q))}.$$

Using the result of Proposition 6.1 we build table $\Phi_1 = \Phi_c(\phi_1, \phi_2, Q)$ on ϕ_1 with depth C_0 and estimate as follows

$$\begin{aligned}
\|\phi_1\phi_2\phi_3\|_{L^p(I^c, C_0(Q))}^p &= \sum_{q'_0 \in \mathcal{Q}_{C_0}(Q)} \|\phi_1\phi_2\phi_3\|_{L^p((1-c)q'_0)}^p \\
&\leq \sum_{q_0, q'_0 \in \mathcal{Q}_{C_0}(Q)} \|\Phi_1^{(q_0)}\phi_2\phi_3\|_{L^p((1-c)q'_0)}^p \\
&= \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} \|\Phi_1^{(q_0)}\phi_2\phi_3\|_{L^p((1-c)q_0)}^p \\
&+ \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} \sum_{q'_0 \in \mathcal{Q}_{C_0}(Q) \setminus \{q_0\}} \|\Phi_1^{(q_0)}\phi_2\phi_3\|_{L^p((1-c)q'_0)}^p \\
&\leq \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} \|\Phi_1^{(q_0)}\phi_2\phi_3\|_{L^p((1-c)q_0)}^p + \left(Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})+\Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i)} \right)^p.
\end{aligned}$$

We owe an explanation on the how we obtained the last term in the sequence of inequalities above. Based on the property (6.4) of tables we conclude that, for each q_0, q'_0 with $q_0 \neq q'_0$ the following hold true

$$\|\Phi_1^{(q_0)}\phi_2\phi_3\|_{L^1((1-c)q'_0)} \leq Cc^{-C} R^{-\frac{n-2}{4}+\Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i)}.$$

Given the size of the cubes, we easily obtain an L^2 estimate of type $\|\phi\|_{L^2(q'_0)} \leq R^{\frac{1}{2}} M(\phi)$ for each term, from which it follows that

$$\|\Phi^{(q_0)}\phi_2\phi_3\|_{L^{\frac{2}{3}}((1-c)q'_0)} \leq CR^{\frac{3}{2}\Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i)}.$$

By interpolation, we obtain the L^p bound

$$\|\Phi_1^{(q_0)}\phi_2\phi_3\|_{L^p((1-c)q'_0)} \leq Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})+\Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i)},$$

which holds true for any $q_0 \neq q'_0$. In conclusion, we have just finished proving the following estimate:

$$\|\phi_1\phi_2\phi_3\|_{L^p(I^c, C_0(Q))}^p \leq \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} \|\Phi_1^{(q_0)}\phi_2\phi_3\|_{L^p((1-c)q_0)}^p + \left(Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})+\Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i)} \right)^p.$$

Next we repeat the procedure for ϕ_2 and ϕ_3 : we construct the tables on ϕ_2, ϕ_3 , $\Phi_2 = \Phi_c(\phi_2, \phi_3, Q)$, $\Phi_3 = \Phi_c(\phi_3, \Phi_1, Q)$ with depth C_0 and estimate in a similar manner (see Remark 1 after Proposition 6.1); we obtain

$$\|\phi_1\phi_2\phi_3\|_{L^p(I^c, C_0(Q))}^p \leq \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} \|\Phi_1^{(q_0)}\Phi_2^{(q_0)}\Phi_3^{(q_0)}\|_{L^p((1-c)q_0)}^p + \left(Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})+\Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i)} \right)^p.$$

We recall that q_0 has size $\frac{4R}{2^{C_0}} \leq \frac{R}{2}$ (which in fact may be seen as setting the threshold needed for C_0). Using this we conclude that

$$\begin{aligned}
\|\phi_1\phi_2\phi_3\|_{L^p(I^{c,C_0}(Q))}^p &\leq \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} \left(\bar{A}_p\left(\frac{R}{2}\right) \right)^p \Pi_{i=1}^3 M(\Phi_i^{(q_0)})^{\frac{p}{2}} \\
&\quad + \left(Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})} \Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i) \right)^p \\
&\leq \left(\bar{A}_p\left(\frac{R}{2}\right) \right)^p \Pi_{i=1}^3 \left(\sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} M(\Phi_i^{(q_0)}) \right)^{\frac{p}{2}} \\
&\quad + \left(Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})} \Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i) \right)^p \\
&\leq \left(\bar{A}_p\left(\frac{R}{2}\right) \right)^p \Pi_{i=1}^3 M(\Phi_i)^{\frac{p}{2}} + \left(Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})} \Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i) \right)^p \\
&\leq (1+cC)^{3p} \left(\bar{A}_p\left(\frac{R}{2}\right) \right)^p \Pi_{i=1}^3 M(\phi_i)^{\frac{p}{2}} + \left(Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})} \Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i) \right)^p.
\end{aligned}$$

where we have used (6.3) in the last line. In using the induction-type bound on $\Pi_{i=1}^3 \Phi_i^{(q_0)}$ we are using the margin bounds on Φ_i from (6.2) to conclude with (3.2); this is easily seen to be the case provided R is large enough to satisfy $CR^{-\frac{1}{2}} \leq R^{-\frac{1}{4}}$. In passing to the second estimate we have used the following sequence inequality

$$\sum_i a_i^q b_i^q c_i^q \leq \left(\sum_i a_i \right)^q \left(\sum_i b_i \right)^q \left(\sum_i c_i \right)^q$$

which holds true for $q \geq \frac{1}{3}$. This easily follows from interpolating the trivial bounds

$$\|abc\|_{l^{\frac{1}{3}}}, \|abc\|_{l^\infty} \leq \|a\|_{l^1} \|b\|_{l^1} \|c\|_{l^1}.$$

where by abc we mean the sequence $\{a_i b_i c_i\}_i$.

Recalling (6.13), we obtain that for any cube Q_R of size R the following holds true

$$\begin{aligned}
\|\phi_1\phi_2\phi_3\|_{L^p(Q_R)} &\leq (1+cC) \Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i) \\
&\quad \cdot \left((1+cC)^{3p} \left(\bar{A}_p\left(\frac{R}{2}\right) \right)^p + \left(Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})} \Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i) \right)^p \right)^{\frac{1}{p}}
\end{aligned}$$

As a consequence we obtain

$$A_p(R) \leq (1+cC) \left((1+cC)^{3p} \left(\bar{A}_p\left(\frac{R}{2}\right) \right)^p + \left(Cc^{-C} R^{\frac{n+4}{2}(\frac{1}{p}-\frac{3}{2}\cdot\frac{n+2}{n+4})} \Pi_{i=1}^3 M^{\frac{1}{2}}(\phi_i) \right)^p \right)^{\frac{1}{p}}.$$

after redefining C , and this is precisely the statement in (3.3). \square

7. THE OPTIMALITY OF $p(k)$

In this section we provide a generalization of the "squashed-cap" construction from [11] with the purpose of establishing the optimality of $p(k)$. To keep things simple, we assume that the surfaces are subsets of the $\mathbb{S}^n \subset \mathbb{R}^{n+1}$, the unit sphere in \mathbb{R}^{n+1} , and moreover that

they are neighborhoods of the points e_i ; here $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ is the unit vector whose coordinates are 1 on the i 'th position and 0 elsewhere.

For each $1 \leq i \leq k$, we construct the rectangular parallelepipeds

$$D_i = \{\zeta \in \mathbb{R}^{n+1}, |\zeta_i - 1| < \epsilon^2, |\zeta_j| < \epsilon^2, j = \{1, \dots, k\} \setminus \{i\}, |\zeta_l| < \epsilon, k+1 \leq l \leq n+1\}.$$

We let $S_i = \mathbb{S}^n \cap D_i$ and note that e_i is transversal to S_i . Thus S_i can be easily parametrized by $\Sigma_i : U_i \subset \{\zeta_i = 0\} \rightarrow S_i$ by

$$\Sigma_i(\zeta_1, \dots, \bar{\zeta}_i, \dots, \zeta_{n+1}) = (\zeta_1, \dots, \zeta_{i-1}, \sqrt{1 - \zeta_1^2 - \dots - \zeta_{i-1}^2 - \zeta_{i+1}^2 - \dots - \zeta_{n+1}^2}, \zeta_{i+1}, \dots, \zeta_{n+1}).$$

Here, the set U_i is characterized by $|\zeta_j| < \epsilon^2, j = \{1, \dots, k\} \setminus \{i\}, \zeta_i = 0, |\zeta_l| < \epsilon, k+1 \leq l \leq n+1$. We pick $f_i = 1$ on U_i , therefore

$$\|f_i\|_{L^2(U_i)} = |U_i|^{\frac{1}{2}} \approx (\epsilon^{2(k-1)} \epsilon^{n+1-k})^{\frac{1}{2}} = \epsilon^{\frac{n+k-1}{2}}$$

where $|U_i|$ is the measure of $U_i \subset \mathbb{R}^n$.

Next we consider the rectangular parallelepiped $R \subset \mathbb{R}^{n+1}$ described by

$$R_c = \{x \in \mathbb{R}^{n+1}, |x_i| \leq c\epsilon^{-2}, i = \{1, \dots, k\}, |x_i| < c\epsilon^{-1}, k+1 \leq i \leq n+1\}.$$

Recall that

$$\mathcal{E}_i f_i(x) = \int_{U_i} e^{ix \cdot \Sigma_i(\xi)} d\xi.$$

By choosing c small enough so that $|x \cdot \Sigma_i(\xi) - x_i| \ll 1$, for all $\xi \in U_i, x \in R_c$, it follows that

$$|\mathcal{E}_i f_i(x)| \gtrsim |U_i| \approx \epsilon^{2(k-1)} \cdot \epsilon^{n+1-k} = \epsilon^{n+k-1}, \forall x \in R_c.$$

Note that the smallness of c above is universal, and, most important, it is independent of ϵ . Then we have

$$\|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p} \gtrsim (\epsilon^{n+k-1})^k |R_c|^{\frac{1}{p}} \approx \epsilon^{k(n+k-1)} \cdot (\epsilon^{-2k} \epsilon^{-(n+1-k)})^{\frac{1}{p}} = \epsilon^{k(n+k-1)} \cdot \epsilon^{-\frac{n+1+k}{p}}$$

Then our results would imply

$$\epsilon^{k(n+k-1)} \cdot \epsilon^{-\frac{n+1+k}{p}} \lesssim \epsilon^{k \cdot \frac{n+k-1}{2}}$$

which, by taking $\epsilon \rightarrow 0$, implies $\frac{k(n+k-1)}{2} - \frac{n+1+k}{p} \geq 0$, thus $p \geq \frac{2(n+k+1)}{k(n+k-1)}$.

The above construction shows the optimality of $p(k)$ in the case of subsets of the sphere and it can be easily adapted to the case of the paraboloid. In fact, it seems to be relevant for hypersurfaces with non-vanishing Gaussian curvature.

But this counterexample works the same way for multi-conical surfaces based on the following simple observation: all that is needed is the consistency of scale ϵ in the $\zeta_{k+1}, \dots, \zeta_{n+1}$ directions with the ϵ^2 scale in the normal direction at each piece; this is a consequence of the quadratic character of the surfaces in the directions $\zeta_{k+1}, \dots, \zeta_{n+1}$. The details are left as an exercise.

REFERENCES

1. Ioan Bejenaru, *The multilinear restriction estimate: a short proof and a refinement*, preprint.
2. ———, *Optimal bilinear restriction estimates for general hypersurfaces and the role of the shape operator*, preprint.
3. Jonathan Bennett, Anthony Carbery, and Terence Tao, *On the multilinear restriction and Kakeya conjectures*, Acta Math. **196** (2006), no. 2, 261–302. MR 2275834 (2007h:42019)

4. J. Bourgain, *Estimates for cone multipliers*, Geometric aspects of functional analysis (Israel, 1992–1994), Oper. Theory Adv. Appl., vol. 77, Birkhäuser, Basel, 1995, pp. 41–60. MR 1353448 (96m:42022)
5. ———, *On the Schrödinger maximal function in higher dimension*, Tr. Mat. Inst. Steklova **280** (2013), no. Ortogonalnye Ryady, Teoriya Priblizhenii i Smezhnye Voprosy, 53–66. MR 3241836
6. Jean Bourgain and Ciprian Demeter, *The proof of the l^2 decoupling conjecture*, arXiv:1403.5335.
7. ———, *The proof of the l^2 decoupling conjecture*, Ann. of Math. (2) **182** (2015), no. 1, 351–389. MR 3374964
8. Jean Bourgain and Larry Guth, *Bounds on oscillatory integral operators based on multilinear estimates*, Geom. Funct. Anal. **21** (2011), no. 6, 1239–1295. MR 2860188 (2012k:42018)
9. Manfredo Perdigão do Carmo, *Riemannian geometry*, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 1992, Translated from the second Portuguese edition by Francis Flaherty. MR 1138207 (92i:53001)
10. Xiumin Du and Xiaochun Li, *Pointwise convergence of the solutions to schrödinger equation in \mathbb{R}^2* , arXiv:1508.05437.
11. Damiano Foschi and Sergiu Klainerman, *Bilinear space-time estimates for homogeneous wave equations*, Ann. Sci. École Norm. Sup. (4) **33** (2000), no. 2, 211–274. MR 1755116 (2001g:35145)
12. Larry Guth, *A restriction estimate using polynomial partitioning*, arXiv:1407.1916.
13. ———, *The endpoint case of the Bennett-Carbery-Tao multilinear Keakey conjecture*, Acta Math. **205** (2010), no. 2, 263–286. MR 2746348 (2012c:42027)
14. Carlos E. Kenig and Frank Merle, *Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case*, Invent. Math. **166** (2006), no. 3, 645–675. MR 2257393 (2007g:35232)
15. Sanghyuk Lee, *Endpoint estimates for the circular maximal function*, Proc. Amer. Math. Soc. **131** (2003), no. 5, 1433–1442 (electronic). MR 1949873 (2003k:42035)
16. ———, *Bilinear restriction estimates for surfaces with curvatures of different signs*, Trans. Amer. Math. Soc. **358** (2006), no. 8, 3511–3533 (electronic). MR 2218987 (2007a:42023)
17. Sanghyuk Lee and Ana Vargas, *Restriction estimates for some surfaces with vanishing curvatures*, J. Funct. Anal. **258** (2010), no. 9, 2884–2909. MR 2595728 (2011j:42029)
18. F. Merle and L. Vega, *Compactness at blow-up time for L^2 solutions of the critical nonlinear Schrödinger equation in 2D*, Internat. Math. Res. Notices (1998), no. 8, 399–425. MR 1628235 (99d:35156)
19. Elias M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. MR 1232192 (95c:42002)
20. T. Tao and A. Vargas, *A bilinear approach to cone multipliers. I. Restriction estimates*, Geom. Funct. Anal. **10** (2000), no. 1, 185–215. MR 1748920 (2002e:42012)
21. ———, *A bilinear approach to cone multipliers. II. Applications*, Geom. Funct. Anal. **10** (2000), no. 1, 216–258. MR 1748921 (2002e:42013)
22. Terence Tao, *Endpoint bilinear restriction theorems for the cone, and some sharp null form estimates*, Math. Z. **238** (2001), no. 2, 215–268. MR 1865417 (2003a:42010)
23. ———, *A sharp bilinear restrictions estimate for paraboloids*, Geom. Funct. Anal. **13** (2003), no. 6, 1359–1384. MR 2033842 (2004m:47111)
24. ———, *Nonlinear dispersive equations*, CBMS Regional Conference Series in Mathematics, vol. 106, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006, Local and global analysis. MR 2233925 (2008i:35211)
25. Thomas Wolff, *A sharp bilinear cone restriction estimate*, Ann. of Math. (2) **153** (2001), no. 3, 661–698. MR 1836285 (2002j:42019)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SAN DIEGO, LA JOLLA, CA 92093-0112 USA

E-mail address: ibejenaru@math.ucsd.edu